

LE ORIGINI DEGLI ETRUSCHI

**UNA CONTROVERSIA MILLENARIA RISOLTA DALLA
GENOMICA**

Silvano Salvador

INTRODUZIONE GENERALE

I. Il fascino degli Etruschi e la struttura del problema

Tra tutte le civiltà dell'Italia antica, quella etrusca occupa una posizione singolare e per molti aspetti paradossale nella memoria culturale europea: è la sola che abbia raggiunto un grado di complessità istituzionale, artistica, tecnica e religiosa pienamente paragonabile a quella greca o romana — con grandi città di cinquantamila abitanti già nell'VIII secolo avanti Cristo, con una rete urbana e commerciale che aveva dominato il Tirreno per secoli, con una disciplina augurale e una teologia dell'espiazione che Roma avrebbe ereditato quasi integralmente — e che al tempo stesso sia rimasta a lungo avvolta in un alone di enigma che le altre grandi civiltà del Mediterraneo antico non dividevano nella stessa misura o nella stessa forma. La grandezza di Roma era patrimonio di un'intera tradizione letteraria in latino leggibile direttamente da chiunque avesse ricevuto un'educazione classica; la grandezza di Atene era accessibile attraverso i testi di Tucidide e Platone, di Sofocle e Aristofane, di Aristotele e Demostene; la grandezza di Cartagine era oscurata ma non silenziosa — vi erano testimonianze latine e greche che ne conservavano almeno il contorno. La grandezza degli Etruschi era leggibile nei loro monumenti — nei templi con le terrecotte policrome di Veio e di Capua, nelle tombe a tumulo di Cerveteri con i loro corredi d'oro lavorato a granulazione finissima, nelle pitture della necropoli di Monterozzi con i loro banchettanti e i loro danzatori, nell'organizzazione idraulica delle grandi città con i loro cunicoli e i loro pozzi —, nella loro metallurgia del bronzo e del ferro che aveva rifornito l'intero bacino occidentale del Mediterraneo dal IX al IV secolo avanti Cristo, nella loro rete di scambi che collegava la Sardegna alla Pianura Padana e l'Etruria alla costa siriana attraverso la mediazione fenicia. Ma questa grandezza non era accessibile dall'interno: non attraverso storie scritte in etrusco da storici etruschi, non attraverso filosofi o poeti etruschi di cui si fossero tramandati i nomi e le opere, non attraverso testi sacri nella lingua originale, non attraverso corrispondenza diplomatica o trattati commerciali. Di una civiltà che aveva prodotto centinaia di migliaia di oggetti, migliaia di tombe, decine di grandi città, secoli di dominio commerciale e politico sull'Italia centrale, ci resta nella lingua originale soltanto un corpus epigrafico di poco più di tredicimila testi, per la quasi totalità di carattere funerario e votivo, formulari al punto da permettere la ricostruzione morfologica della lingua ma non quella del suo lessico in misura sufficiente a una comprensione piena.

Questa asimmetria fondamentale — tra la grandezza monumentale e la quasi totale assenza di una voce narrativa in lingua etrusca — ha alimentato per secoli un fascino che si è spesso trasformato in proiezione ideologica. Gli Etruschi sono diventati, nel corso del tempo, lo specchio su cui generazioni di lettori, di studiosi e di appassionati hanno riflesso le proprie domande sull'identità,

sull'origine, sull'alterità. Sono stati di volta in volta gli antenati dei Toscani, i portatori della civiltà orientale in Occidente, i rappresentanti di un'Europa pre-indoeuropea perduta, i testimoni di un'età dell'oro precristiana, le vittime di una colonizzazione culturale romana che avrebbe soppresso la loro voce. Nessuna di queste proiezioni è interamente falsa — la civiltà etrusca è stata tutte queste cose per qualcuno, in qualche momento storico — ma nessuna è interamente vera, e tutte hanno distorto la ricerca sulle origini in modi che sono analizzabili e identificabili con gli strumenti della storia della storiografia.

La questione delle origini è stata il centro magnetico di questo fascino perché concentra in modo emblematico tutte le tensioni del problema etrusco. Da dove venivano? Erano italiani nel senso di autoctoni, figli della stessa terra che avevano civilizzato? Erano emigrati dall'Oriente, portatori di un'antichità ancora più remota che l'Italia aveva ricevuto come dono dal Mediterraneo orientale? Oppure erano scesi dalle Alpi, depositari di una tradizione nordica che Roma aveva poi assorbito e trasformato? Queste domande, che il pensiero antico aveva già formulato con notevole lucidità nel V secolo avanti Cristo — Erodoto, Dionisio di Alicarnasso, Tito Livio, tre risposte diverse che esaurivano lo spazio logico delle possibilità geografiche principali — non hanno ricevuto risposta definitiva per venticinque secoli. Non per mancanza di interesse, non per assenza di intelletti acuti, ma perché gli strumenti necessari a produrre evidenza discriminante non esistevano. Questa monografia racconta la lunga attesa e la sua risoluzione.

Per comprendere la natura di quella lunga attesa è necessario capire che cosa significa, in termini epistemologici, avere una domanda senza gli strumenti per rispondervi. La domanda sulle origini degli Etruschi era, nella sua dimensione demografica e biologica fondamentale, una domanda sulla storia delle popolazioni: su chi abitasse un certo territorio in un certo momento, da dove fosse arrivato, con quale popolazione precedente fosse in relazione di continuità o di discontinuità. Questo tipo di domanda richiede, per essere risposto in modo non ambiguo, un tipo di evidenza che sia al tempo stesso diretta — non mediata da inferenze culturali o materiali che possono essere interpretate in modi diversi — e quantitativamente precisa, con barre di errore esplicite e con potere discriminante misurabile. La filologia classica, l'archeologia stratigrafica, la linguistica storica comparativa, l'antropologia fisica craniometrica — tutti gli strumenti che hanno affrontato la questione tra il Settecento e il 2018 — erano incapaci di produrre evidenza con queste caratteristiche. Non erano strumenti scadenti in assoluto: erano strumenti scadenti per questa specifica domanda, nel senso tecnico e preciso che la struttura logica di ciascuno di essi non permetteva di produrre rapporti di verosimiglianza sufficientemente lontani da 1 rispetto alle ipotesi alternative. La paleogenomica è il primo strumento che permette di produrre tale evidenza, e questa

è la ragione — non la qualità superiore degli studiosi che la praticano — per cui ha risolto in pochi anni una controversia che aveva resistito per venticinque secoli.

II. Il corpus epigrafico etrusco: dimensioni, distribuzione e limiti strutturali

Prima di addentrarsi nella questione delle origini nella sua complessità multidimensionale, è indispensabile fornire al lettore una mappa precisa di ciò che sappiamo della lingua etrusca attraverso il suo corpus epigrafico — poiché è da questa mappa che derivano molte delle domande irrisolvibili con gli strumenti della filologia classica e molte delle qualificazioni necessarie alla lettura dei risultati genomici. La lingua etrusca è nota esclusivamente attraverso le iscrizioni — non vi sono manoscritti etruschi, non vi sono codici medievali, non vi sono trasmissioni indirette in traduzione, non vi sono grammatiche antiche composte da autori etruschi o da grammatici romani interessati alla lingua. La tradizione latina che conserva il ricordo della letteratura etrusca — i libri etruschi che includevano i Libri Acherontici, i Libri Fatales, i Libri Rituales, i Libri Haruspici, i Libri Fulgurales — non ci ha trasmesso questi testi in originale: ci ha trasmesso solo il ricordo del loro contenuto attraverso la mediazione di autori latini come Cicerone, Livio, Plinio il Vecchio e Seneca, che descrivevano la disciplina etrusca senza citarne i testi originali se non in rarissimi frammenti.

Il corpus epigrafico noto comprende oggi approssimativamente tredicimila testi, con stime più recenti che arrivano fino a quindicimila quando si includono i frammenti di leggibilità incerta e le iscrizioni di recentissima scoperta non ancora entrate nei repertori ufficiali. La grande maggioranza di questi testi — nell'ordine dell'ottantacinque per cento — è di carattere funerario: epitaffi, didascalie tombali, formule di dedicazione della tomba, indicazioni del nome del defunto con la sua filiazione e la sua età. Una quota significativa — intorno al dieci per cento — è di carattere votivo: dediche a divinità su oggetti bronzei, su ceramiche, su lamine di piombo deposte nei santuari. Il rimanente cinque per cento comprende testi di natura più complessa e quindi più preziosi per la comprensione della lingua: testi commerciali, testi con contenuto numerale e calendárico, testi di incerta funzione, testi bilingui. Questa distribuzione tipologica è il dato più importante per comprendere i limiti strutturali della nostra conoscenza dell'etrusco: i testi funerari e votivi sono prevalentemente formulari — le stesse strutture, gli stessi sintagmi, le stesse sequenze nominali e verbali si ripetono con variazioni minime in migliaia di iscrizioni — il che permette di ricostruire con grande precisione la morfologia flessiva della lingua ma offre pochissimo materiale lessicale non onomastico. Il lessico etrusco ricostruibile con certezza, al netto dei nomi propri, non supera le due-trecento radici — una cifra derisoria rispetto alle necessità di qualsiasi analisi linguistica comparativa sistematica che voglia andare oltre il livello morfologico.

I testi di maggiore importanza per la comprensione della lingua etrusca nelle sue dimensioni non formulari sono un gruppo ristretto di documenti eccezionali. Il più lungo testo etrusco noto è la cosiddetta *Liber Linteus Zagrabensis* — il Libro di lino di Zagabria — un testo scritto su bende di lino che, in un momento non precisamente determinato ma probabilmente nel II–I secolo avanti Cristo, fu riutilizzato per avvolgere una mummia egiziana di sesso femminile. La mummia fu importata in Europa nel XIX secolo da un funzionario asburgico, Mihajlo Barić, che la acquistò in Egitto intorno al 1848 e la portò a Vienna, da dove fu successivamente trasferita al Museo Archeologico di Zagabria. Il contenuto del testo rimase ignoto fino al 1891, quando lo studioso tedesco Heinrich Brugsch riconobbe i caratteri come etruschi anziché egiziani. Il testo originale, che occupava le dodici strisce di lino prima del loro riutilizzo come bende, è databile su base paleografica e contenutistica al III–II secolo avanti Cristo; è un calendario sacro o un manuale rituale che contiene formule ripetute riferibili a cerimonie da compiersi in determinati giorni del calendario, con ricorrenti riferimenti a divinità etrusche come Uni, Nethuns, Catha e Thalna. Il testo comprende circa milleduecento parole leggibili — molte altre sono distrutte o illeggibili per lo stato di conservazione delle bende — di cui un numero significativo ricorre più volte nelle dodici colonne, riducendo il numero di lemmi distinti a circa trecento. Nonostante la sua eccezionale lunghezza rispetto alla media del corpus, il *Liber Linteus* rimane largamente opaco all'interpretazione semantica precisa per assenza di un testo bilingue di riferimento: sappiamo che certi termini si riferiscono a divinità o a rituali perché ricorrono in contesti analoghi ad altri testi etruschi, non perché il significato sia dimostrabile indipendentemente.

Un secondo documento di straordinaria importanza è la *Tabula Cortonensis*, una lamina di bronzo composta da otto frammenti, scoperta a Cortona nel 1992 durante lavori di costruzione e pubblicata criticamente nel 1999 da Lucio Agostiniani e Francesco Nicosia. La lamina, databile paleograficamente al II secolo avanti Cristo, contiene un testo di trecentosettantadue parole disposto su otto righe del recto e una riga del verso. Il testo è con ogni probabilità una transazione giuridica relativa alla compravendita di un terreno o alla regolamentazione di diritti fondiari: nomina diversi individui con i loro nomi completi — prenome, gentilizio, cognomen e patronimico — indica localizzazioni geografiche identificabili con luoghi del Cortonese, e include formule che sembrano riferirsi a garanzie legali, testimonianze e obbligazioni. La *Tabula Cortonensis* è uno dei rarissimi testi etruschi con contenuto non funerario e non votivo che permette di inferire strutture sintattiche più complesse delle formule onomastiche standard, e la sua pubblicazione ha permesso di identificare alcune strutture della sintassi della frase etrusca che non erano precedentemente documentate. Anche questo testo, tuttavia, rimane parzialmente opaco: la terminologia giuridica

specifica — nomi dei contratti, dei diritti, delle procedure — è difficilmente interpretabile senza un testo bilingue di riferimento.

Le *Laminette d'oro di Pyrgi*, scoperte nel 1964 durante gli scavi del santuario costiero di Santa Severa — l'antica Pyrgi, porto di Cerveteri — e databili con buona approssimazione al 500–480 avanti Cristo, costituiscono il documento bilingue più importante della tradizione epigrafica etrusca e il punto di ancoraggio lessicale più solido disponibile per la lingua. Le tre laminette, conservate al Museo Nazionale Etrusco di Villa Giulia a Roma, recano due testi in etrusco e uno in fenicio punico — non traduzioni letterali l'uno dell'altro, ma testi che registrano la stessa dedicazione del santuario alla dea Uni (identificata con Astarte nella versione fenicia) con formulazioni linguisticamente indipendenti. Il confronto tra il testo etrusco e quello fenicio ha permesso di identificare con certezza il significato di alcuni termini etruschi — in particolare il verbo *vacl* che indica la donazione o dedicazione rituale, e alcuni titoli religiosi — confermando interpretazioni già proposte per via interna al corpus e fornendo un punto di ancoraggio verificabile per l'interpretazione della sintassi della frase nominale etrusca.

Il *Cippo Perugino*, un cippo di travertino con iscrizione di centoventidue righe scoperto a Perugia nel 1822 durante la costruzione di una cisterna e conservato al Museo Nazionale di Perugia, è il secondo testo più lungo del corpus dopo il Liber Linteus. È con ogni probabilità un accordo di confine o un contratto fondiario tra due famiglie — la famiglia Velthina e la famiglia Afuna — che regola i diritti sull'uso di un fondo. La struttura del testo è relativamente chiara: prevede una descrizione della situazione esistente, un accordo sulle condizioni future, e una formula di sanzione per il mancato rispetto dell'accordo. Anche in questo caso, tuttavia, la terminologia giuridica specifica rimane in larga parte opaca.

A questi grandi documenti si aggiungono alcune centinaia di iscrizioni di lunghezza intermedia — tra trenta e duecento parole — che provengono principalmente dai santuari (come le lamine di piombo con dediche votive di Punta della Vipera, le lamine bronzee di Magliano, le iscrizioni del santuario di Gravisca) e che offrono materiale linguistico più vario dei semplici epitaffi. Le lamine di piombo con iscrizioni sono particolarmente preziose perché il piombo è un supporto deperibile ma impermeabile, conservato in condizioni anaerobiche nei depositi votivi interrati, e i testi vi si trovano spesso in buono stato di leggibilità nonostante l'età. Alcune di queste lamine contengono testi di contenuto magico o di defixio — testi di maledizione contro specifici individui — che costituiscono un genere documentato anche in greco e in latino e che permette qualche confronto tipologico utile per l'interpretazione lessicale.

III. La paleografia etrusca: dall'abecedario di Marsiliana alla fase recente

L'etrusco si scrive in un alfabeto derivato dall'alfabeto greco per via di trasmissione diretta, e la storia di questa derivazione è uno dei contributi più importanti della paleografia etrusca alla comprensione della civiltà nel suo complesso. L'adattamento dell'alfabeto greco all'etrusco non avvenne in modo astratto o attraverso una mediazione indiretta: avvenne attraverso il contatto diretto tra popolazioni etrusche e colonie greche dell'Italia meridionale e della costa tirrenica nel corso dell'VIII–VII secolo avanti Cristo, in un momento storicamente circoscritto e identificabile con ragionevole precisione attraverso i dati paleografici comparativi. Questo fatto — la dipendenza dell'alfabeto etrusco da un modello greco euboico specifico — è uno degli elementi che legano la storia della civiltà etrusca alla storia dei commerci mediterranei nel periodo orientalizzante.

L'alfabeto greco che gli Etruschi adottarono come modello appartiene al tipo euboico — il tipo in uso nelle colonie greche di Pithecusa, fondata intorno all'770 avanti Cristo sull'isola di Ischia, e di Cuma, fondata sulla terraferma campana poco dopo. Questa identificazione è documentata da una serie di caratteristiche specifiche dell'alfabeto etrusco che lo distinguono dagli alfabeti greci di tipo ionico-attico o corinzio: in particolare, la presenza della lettera *waw* (il digamma greco F, che rappresenta il suono /w/ e che i dialetti greci ionici avevano già perduto o stavano perdendo nel momento in cui gli Etruschi adottarono l'alfabeto), la forma caratteristica del sigma a quattro tratti (vs. il sigma attico Σ), e la presenza del *qoppa* (Q, una lettera dell'alfabeto semitico che i Greci avevano adottato per distinguere la velare /k/ davanti a vocali posteriori e che i dialetti ionici stavano eliminando, ma che l'alfabeto euboico conservava ancora). Il vettore della trasmissione era quasi certamente il commercio: i contatti tra il centro etrusco di Cerveteri e la colonia greca di Pithecusa sono documentati dai ritrovamenti ceramici, in particolare di anfore etrusche nei depositi di Pithecusa e di ceramica greca euboica nei depositi di Cerveteri, già nell'VIII secolo avanti Cristo.

L'alfabeto etrusco più antico — documentato dall'*abecedario di Marsiliana d'Albegna*, una tavoletta d'avorio con l'alfabeto completo inciso sul bordo datata al 700–675 avanti Cristo sulla base del contesto stratigrafico e del confronto con ceramica greca protocorinzia, conservata al Museo Nazionale di Firenze — contiene tutte le ventisei lettere dell'alfabeto greco euboico, incluse alcune che l'etrusco non utilizzò mai nella scrittura effettiva perché il sistema fonologico della lingua non richiedeva i suoni corrispondenti. La presenza di queste lettere «inutilizzate» nell'abecedario — in particolare il *beta* (B) per il suono /b/ che l'etrusco non distingueva dal /p/, il *delta* (D) per il suono /d/ assente nel sistema fonemico etrusco, il *omicron* (O) per una distinzione vocalica che l'etrusco non manteneva come contrasto fonologico, e alcune varianti di sibilante — documenta in modo inequivocabile che la trasmissione avvenne per via di copia diretta di un modello greco completo, non attraverso un adattamento già selettivo delle sole lettere funzionalmente necessarie. Questa

osservazione è metodologicamente importante: se l'alfabeto fosse stato inventato indipendentemente o trasmesso attraverso una mediazione culturale indiretta, ci aspetteremmo un sistema già adattato alle necessità fonologiche dell'etrusco. La presenza delle lettere inutilizzate dimostra invece che qualcuno — probabilmente un artigiano o un mercante di Cerveteri o di Vulci — copiò un abecedario greco euboico integrale e lo usò come modello per addestrare gli scribi etruschi, che poi selezionarono le lettere utili e scartarono le altre nella pratica scrittoria.

La storia dell'alfabeto etrusco è divisibile in tre grandi fasi cronologiche, ciascuna caratterizzata da sviluppi paleografici identificabili che permettono di datare le iscrizioni in assenza di contesti stratigrafici certi. La **fase arcaica** (700–500 avanti Cristo) è caratterizzata dall'uso di forme grafiche che mantengono la massima fedeltà al modello greco euboico: le lettere hanno forme angolate, il ductus è spesso irregolare con lettere di dimensioni variabili, la direzione della scrittura non è ancora stabilizzata (si trovano iscrizioni bistrofediche, cioè con direzione alternata da sinistra a destra e da destra a sinistra, accanto a iscrizioni interamente da destra a sinistra o interamente da sinistra a destra). La **fase classica** (500–300 avanti Cristo) vede la stabilizzazione delle forme grafiche proprie dell'etrusco, con la scelta definitiva della direzione da destra a sinistra, la normalizzazione delle forme delle lettere, l'eliminazione delle lettere inutilizzate, e uno ductus più fluido e più uniforme tra i diversi centri. È in questa fase che si sviluppano le varianti regionali dell'alfabeto etrusco, con differenze documentabili tra l'area meridionale (Cerveteri, Tarquinia, Vulci, Veio) e l'area settentrionale (Chiusi, Arezzo, Cortona, Perugia, Bologna). La **fase recente** (300–50 avanti Cristo) è caratterizzata da ulteriori semplificazioni grafiche — riduzione del numero di segni distinti, assimilazione di alcuni segni — e da fenomeni di convergenza con l'alfabeto latino che riflettono il progressivo bilinguismo etrusco-latino della popolazione etrusca in seguito all'espansione romana. L'ultima iscrizione etrusca datata con sicurezza risale al 50–40 avanti Cristo, e da questo momento l'etrusco sopravvive al più come lingua liturgica in contesti rituali molto circoscritti.

Un aspetto paleografico di grande rilevanza per la questione delle origini è la **variazione regionale dell'alfabeto etrusco**: non esiste un unico alfabeto etrusco uniforme valido per tutta l'Etruria, ma una famiglia di varianti regionali che si distribuiscono geograficamente in modo non casuale e che riflettono storie di contatto e di istruzione diverse. Le differenze più significative riguardano il trattamento delle sibilanti — l'etrusco settentrionale usa forme diverse da quello meridionale per il sigma e lo zeta — e la forma di alcune occlusive. Questo dato è rilevante per la questione delle origini perché esclude la possibilità che l'alfabeto etrusco sia arrivato nell'area come sistema già standardizzato proveniente da un'unica fonte esterna: la variazione regionale è compatibile solo con una diffusione interna graduale a partire da pochi centri di adozione iniziale, probabilmente Cerveteri e Vulci nell'area meridionale, con successiva irradiazione verso l'interno e verso nord.

IV. La struttura di questa monografia: architettura dell'argomentazione

Questa monografia è organizzata in quattro capitoli principali seguiti da un blocco di appendici, una serie di tavole sinottiche e una bibliografia ragionata. L'organizzazione riflette una progressione che va dalla mappa del problema (Capitolo I) all'analisi dell'evidenza risolutiva (Capitolo II), alla ricostruzione storica del dibattito (Capitolo III), alle conclusioni epistemologiche e all'agenda di ricerca futura (Capitolo IV).

Il **Capitolo I** costruisce la mappa epistemologica del problema nella sua complessità multidimensionale. Questo significa, in primo luogo, analizzare le tre ipotesi classiche — anatolica, autoctona, alpina — con il rigore filologico necessario: esaminare i testi antichi in originale, valutarne il potere probatorio con gli strumenti dell'epistemologia storica, identificarne i presupposti metodologici impliciti. Significa, in secondo luogo, affrontare con la profondità tecnica necessaria la questione della lingua etrusca nella sua complessità linguistica: il suo sistema fonologico, la sua morfologia, il problema dell'isolamento genetico, il confronto lemnico-etrusco e le sue implicazioni per le diverse ipotesi di origine. Significa, in terzo luogo, costruire una mappa dei quattro strati analitici del problema — la domanda biologica, la domanda dell'ascendenza stratificata, la domanda cronologica, la domanda della corrispondenza tra profilo genomico e distribuzione culturale — che sono logicamente indipendenti e richiedono metodi diversi per essere affrontati. Significa, infine, elaborare una stima bayesiana formale della probabilità a priori delle ipotesi — cioè la probabilità da assegnare a ciascuna sulla base di tutte le informazioni disponibili prima degli studi paleogenomici — che rende esplicita la struttura logica dell'aggiornamento che seguirà.

Il **Capitolo II** presenta e discute l'analisi paleogenomica degli anni 2019–2021 con il grado di dettaglio tecnico che una monografia accademica richiede. Questo significa descrivere il campione nella sua costruzione e nei suoi limiti, spiegare i metodi di analisi genomica — PCA, ADMIXTURE, qpAdm, qpGraph — nella loro struttura formale e nelle loro assunzioni, presentare i risultati nella loro stratificazione di certezze e probabilità, e produrre l'aggiornamento bayesiano formale che incorpora i risultati nel framework costruito nel Capitolo I. Questo capitolo include anche una sezione sui limiti metodologici della paleogenomica — i rischi di contaminazione, la dipendenza dal pannello di riferimento, il limite della potenza statistica per componenti di piccola dimensione — che sono necessari per una lettura critica e non ingenua dei risultati.

Il **Capitolo III** ripercorre la storiografia della questione dal Settecento al 2019, analizzando i meccanismi istituzionali e cognitivi che hanno mantenuto la controversia artificialmente aperta. Questo non è un capitolo di storia antiquaria della disciplina: è un'analisi sociologica ed epistemologica di come si formano e si mantengono i paradigmi in una disciplina con determinate

strutture istituzionali, con determinate pressioni identitarie e nazionalistiche, e con determinati incentivi accademici. Il caso etrusco è un laboratorio eccezionale per questa analisi perché è abbastanza vicino nel tempo da permettere l'accesso agli archivi e alle corrispondenze private, e abbastanza lontano da permettere una valutazione distaccata.

Il **Capitolo IV** trae le conclusioni epistemologiche, formula con precisione le domande che rimangono aperte dopo la risoluzione genomica e propone un'agenda di ricerca articolata in priorità di breve, medio e lungo termine per i decenni futuri. In particolare, la sezione IV.2 include una sottosezione (IV.2.4) dedicata all'applicazione dei metodi della linguistica computazionale al problema della cronologia e della posizione genetica dell'etrusco-lemnico, e una sezione IV.5 che articola un'agenda integrata di ricerca linguistica su tre orizzonti temporali — breve (infrastruttura digitale del corpus), medio (glottocronologia bayesiana e network analysis delle isoglosse egeo-tirreniche) e lungo (grammatica comparata del pre-indoeuropeo mediterraneo). La domanda sulla lingua rimane il principale problema irrisolto degli studi etruschi. Le appendici offrono strumenti di approfondimento tecnico in otto direzioni, selezionate per stretta pertinenza all'argomentazione centrale: la cronologia ragionata della civiltà etrusca (App. A), la descrizione dei principali siti del campione paleogenomico (App. B), la documentazione filologica completa del confronto lemnico-etrusco (App. C), la guida tecnica alla metodologia paleogenomica (App. D), le tavole sinottiche riassuntive (App. E), la storia e la critica dell'antropologia fisica craniometrica (App. H), i fondamenti matematici delle statistiche-f (App. M), la cronologia degli studi paleogenomici sugli Etruschi 2013–2024 (App. O) e i casi comparativi internazionali (App. P). Le appendici a carattere enciclopedico generale sulla civiltà etrusca — religione, città, commerci, sistema sociale, relazioni con le altre civiltà italiche — non sono incluse in questo volume per ragioni di coerenza tematica: appartengono a un'opera di respiro diverso, più ampio. Il lettore interessato troverà i riferimenti fondamentali nella sezione III della Bibliografia.

Una nota sul tono e sull'approccio. Questa monografia è stata scritta con la convinzione che la precisione tecnica non sia incompatibile con la chiarezza espositiva, ma che sia invece una sua preconditione. Semplificare al punto di distorcere — presentare come certa un'evidenza che è probabilistica, o come risolta una questione che rimane genuinamente aperta, o come univoca un'interpretazione che è controversa — sarebbe un atto di disonestà intellettuale verso il lettore, particolarmente grave in un campo in cui le implicazioni identitarie hanno storicamente condizionato la ricezione della ricerca. L'obiettivo è quindi essere precisi nelle affermazioni, trasparenti nelle incertezze, espliciti nelle strutture logiche degli argomenti, e onesti nell'indicare i limiti di ciò che l'evidenza disponibile permette di concludere.

CAPITOLO ILA DOMANDA E I SUOI STRUMENTI: MAPPA EPISTEMOLOGICA DEL PROBLEMA

I.1 Una domanda che la filologia non ha potuto risolvere: struttura dell'asimmetria epistemica

Alcune domande dell'antichità resistono non perché gli storici abbiano mancato di porle con la giusta intelligenza, ma perché le pongono con strumenti strutturalmente inadeguati alla loro natura. La questione dell'origine degli Etruschi appartiene a questa categoria con una precisione quasi esemplare: è stata formulata con notevole chiarezza già nel V secolo avanti Cristo, è rimasta oggetto di dibattito aperto e inconcludente per venticinque secoli, e ha trovato risposta definitiva soltanto quando il tipo di evidenza necessaria — dati genomici di alta copertura estratti da resti scheletrici adeguatamente datati con radiocarbonio AMS — è diventato tecnicamente accessibile grazie alla rivoluzione della paleogenomica degli anni 2010. La persistenza della controversia non è un'anomalia che si possa spiegare con la negligenza degli studiosi o con il pregiudizio intellettuale dei singoli — benché entrambi abbiano avuto un ruolo in momenti specifici della tradizione storiografica, come si mostrerà nel Capitolo III. È piuttosto la conseguenza strutturale di un'asimmetria epistemica profonda e, in retrospettiva, perfettamente riconoscibile: le fonti disponibili per secoli erano in linea di principio incapaci di discriminare tra le ipotesi in competizione con la risoluzione cronologica e demografica richiesta. Questa incapacità non era un limite della bravura dei ricercatori: era un limite degli strumenti stessi, nel senso tecnico e preciso che la struttura logica degli strumenti disponibili non permetteva di produrre rapporti di verosimiglianza sufficientemente lontani da 1 rispetto alle ipotesi alternative.

Per comprendere perché questa asimmetria fosse strutturalmente insuperabile con gli strumenti tradizionali, è necessario analizzare con precisione quali domande ciascuno strumento era in linea di principio in grado di rispondere. La **filologia classica** poteva ricostruire e interpretare le fonti letterarie antiche sulla questione, ma non poteva produrre evidenza indipendente da quelle fonti per valutarne la plausibilità: la filologia lavora all'interno del canone testuale, non al di fuori di esso. L'**archeologia stratigrafica** poteva documentare la continuità o la discontinuità del record materiale in determinati siti, ma la continuità stratigrafica della cultura materiale — la presenza degli stessi tipi ceramici, degli stessi riti funerari, degli stessi stili decorativi in sequenza ininterrotta — è compatibile con scenari demografici molto diversi. Una popolazione può mantenere le sue tradizioni materiali pur essendo parzialmente sostituita da immigranti che le adottano; e viceversa, una popolazione

biologicamente continua può cambiare radicalmente le sue tradizioni materiali per effetto di cambiamenti religiosi, economici o politici interni. La **linguistica storica comparativa** poteva identificare parentele e divergenze linguistiche tra l'etrusco e le lingue note, ma non poteva determinare se una parentela riflettesse migrazione di popolazione o diffusione di lingua per contatto culturale senza spostamento demografico — una distinzione fondamentale che la linguistica storica, da sola, non è in grado di risolvere. L'**antropologia fisica craniometrica**, infine, misurava caratteri metrici craniali e post-craniali che sono influenzati da fattori ambientali e epigenetici al punto da rendere inaffidabile qualsiasi inferenza demografica a lungo raggio — come si mostrerà nel dettaglio nel Capitolo II. Soltanto la **paleogenomica** — la disciplina che estrae e analizza DNA genomico da resti scheletrici antichi — poteva rispondere direttamente alla domanda biologica-demografica con la risoluzione necessaria, perché lavora su oggetti — i marcatori genomici di singolo nucleotide, gli SNP — la cui struttura causale è ben compresa dai modelli della genetica delle popolazioni e che portano un segnale demografico quantitativamente interpretabile con barre di errore esplicite.

La struttura dell'asimmetria epistemica che ha mantenuto la controversia aperta per venticinque secoli può essere descritta in modo formale attraverso il concetto di *rapporto di verosimiglianza*. Data un'evidenza E e due ipotesi in competizione H_1 e H_2 , il rapporto di verosimiglianza è il rapporto $P(E|H_1)/P(E|H_2)$ — cioè il rapporto tra la probabilità di osservare l'evidenza E se H_1 è vera e la probabilità di osservarla se H_2 è vera. Un rapporto di verosimiglianza vicino a 1 indica un'evidenza non discriminante: quella evidenza è altrettanto probabile sotto entrambe le ipotesi, e il suo osservazione non sposta in modo significativo la probabilità relativa delle due ipotesi. Un rapporto molto lontano da 1 — nell'ordine di 10, 100, 1000 o superiore — indica un'evidenza altamente discriminante che favorisce fortemente una delle due ipotesi rispetto all'altra. Tutta la storia del dibattito sulle origini etrusche prima del 2019 è la storia di evidenze con rapporti di verosimiglianza vicini a 1 rispetto alle ipotesi in competizione: la continuità stratigrafica, la critica linguistica all'ipotesi lidia, i confronti craniometrici — tutte le linee di evidenza tradizionali erano compatibili, con opportuni aggiustamenti, tanto con l'ipotesi autoctona quanto con l'ipotesi anatolica. La paleogenomica ha prodotto evidenze con rapporti di verosimiglianza nell'ordine di 10^2 – 10^3 , e questa differenza di ordini di grandezza — non la qualità degli studiosi o la loro apertura mentale — è la ragione per cui la risoluzione è arrivata così rapidamente dopo l'applicazione del nuovo strumento.

I.2 Le tre ipotesi classiche: analisi filologica delle fonti, struttura logica e potere argomentativo

1.2.1 L'ipotesi anatolica: Erodoto, la tradizione lidia e l'analisi critica del passo

La prima ipotesi, quella dell'origine anatolica o lidia, è associata al nome di Erodoto e al celebre passo del primo libro delle *Storie* (I, 94) che merita di essere letto con la massima attenzione filologica possibile, perché ogni parola di questo testo ha avuto un peso straordinario nella storia del dibattito e perché la sua corretta interpretazione critica — spesso trascurata dalla tradizione storiografica che lo citava senza discuterlo — è fondamentale per valutarne il contenuto probatorio effettivo. Il passo segue immediatamente una descrizione delle consuetudini dei Lidi che Erodoto ha appena confrontato con quelle dei Greci, e si inserisce in un contesto più ampio di excursus etnografici con cui lo storico alterna la narrazione dei fatti politici.

Λέγουσι δὲ σφέας Λυδοὶ [...] ἐπὶ Ἄττος τοῦ Μάνεω βασιλέος σιτοδείην ἰσχυρὴν ἀνὰ τὴν Λυδίην γενέσθαι, καὶ τοὺς Λυδοὺς χρόνον μὲν τινα διαγεῖν, μετὰ δέ, ὥς οὐκ ἀνιέναι, ἀκεόμενοι δὲ τὰ μὲν, τὰ δὲ οὐ, ζητέειν δόξαντα σφί ἐκπέμψαι τε καὶ ἐν τῇ Ὀμβρική κτίσαντας πόλιας οἰκέειν ἐκεῖ. Τυρσηγόν τε ἐπὶ τοῦτοις ἡγεμόνα ἐπιστήσαντες, τοὺς δὲ λειπομένους αὐτοῦ μεταβαλεῖν τε τὸ σῶνμα ἐπὶ τοῦ παιδὸς τοῦ ἡγησαμένου σφέων, καλέεσθαι Τυρσηγούς.

Il greco del passo è notevolmente preciso sulle formule epistemologiche usate. Il verbo *λέγουσι* — «dicono» — all'inizio del brano segnala che la notizia che segue non è di Erodoto ma della tradizione orale lidia. Erodoto non dice «io so» o «è avvenuto» ma «i Lidi dicono che»: è una formula di distanziamento epistemico che nella pratica storiografica erodotea corrisponde all'attribuzione di una notizia a una fonte non verificata direttamente. Questo distanziamento è sistematico nella pratica storiografica di Erodoto, che distingue accuratamente tra ciò che ha visto di persona (*εἶδον αὐτός*), ciò che ha sentito da testimoni diretti (*ἤκουσα*) e ciò che la tradizione narra (*λέγεται* o *λέγουσι*). La narrazione dell'origine lidia degli Etruschi appartiene alla terza categoria — la categoria di qualità probatoria più bassa nella gerarchia epistemica erodotea.

La narrazione ha la struttura inconfondibile di un *aition* — un racconto eziologico che spiega un nome attraverso un evento fondativo nel passato mitico. La struttura è: (1) evento scatenante (la carestia), (2) decisione collettiva (divisione del popolo in due metà), (3) scelta divina o casuale (la sorte), (4) partenza del gruppo, (5) fondazione in un nuovo territorio, (6) adozione del nome dell'eponimo. Questa struttura narrativa è un topos riconoscibile nella tradizione letteraria greca: il racconto dell'origine di un nome attraverso la discendenza da un eroe eponimo — Ione per gli Ioni, Doron per i Dori, Tirreno per i Tirreni — è una forma narrativa convenzionale che non ha di per sé alcuna pretesa di accuratezza storica. Il fatto che la carestia scatenante abbia luogo «al tempo del re Atis figlio di Mane» aggiunge un dettaglio di pseudo-precisione cronologica — nomi di re lidii storicamente documentati — ma questo dettaglio non migliora la qualità della narrazione come evidenza storica: inserisce la migrazione in un quadro cronologico vago («epoca lontana») attraverso

riferimenti a una genealogia regale lidia che Erodoto riporta altrove senza pretendere di verificarne la storicità.

Un secondo aspetto filologico fondamentale è l'identificazione dei Tirreni erodotei con gli Etruschi della tradizione latina e italiana. I Greci usavano abitualmente il termine *Τυρσηνοί* o *Τυρρηνοί* per designare gli Etruschi, e questo nome è quello che Erodoto usa nel passo in questione. L'etimologia del termine era oggetto di dibattito già nell'antichità: Dionisio di Alicarnasso la rigetta come non etrusca, e la questione rimane aperta nella linguistica moderna. La derivazione erodotea da un eponimo Tirreno (*Τυρσηνός*) è chiaramente un'etimologia popolare retrodatata: il nome greco *Τυρσηνοί* esisteva prima che Erodoto scrivesse, e la narrazione dell'eponimo Tirreno è costruita a partire dal nome, non viceversa. Questo tipo di etimologia popolare — spiegare un nome attraverso un antenato omonimo — è uno dei meccanismi narrativi più comuni nella tradizione greca e non ha alcun valore come evidenza storica indipendente.

Nonostante questi limiti probatori evidenti a qualsiasi lettore filologicamente avvertito, il passo erodoteo ha goduto per secoli di un'autorità sproporzionata alla sua qualità di testimonianza storica. Il fattore principale di questa autorità è la posizione di Erodoto nel canone: è il primo autore superstite a trattare la questione in modo esteso, il che lo ha collocato automaticamente nella posizione di punto di partenza obbligato del dibattito. Chiunque scrivesse dopo di lui era costretto a posizionarsi rispetto a lui: chi condivideva l'ipotesi anatolica citava Erodoto come autorità a sostegno; chi la rifiutava doveva prima confutare Erodoto e poi proporre un'alternativa. In entrambi i casi, Erodoto era sempre al centro del dibattito anche quando veniva criticato — un effetto di ancoraggio cognitivo che ha deformato la percezione del peso dell'evidenza per oltre due millenni.

1.2.2 L'ipotesi autoctona: Dionisio di Alicarnasso e il metodo comparativo

La seconda ipotesi, quella dell'origine autoctona, è associata al nome di Dionisio di Alicarnasso e all'argomentazione sviluppata nel primo libro delle *Antichità Romane* (I, 26–30), scritto in greco probabilmente intorno al 7 avanti Cristo durante il soggiorno romano dell'autore, uno storico e retore greco di Alicarnasso trasferitosi a Roma intorno al 30 avanti Cristo e lì rimasto per il resto della sua vita. Dionisio è, tra i critici antichi dell'ipotesi erodotea, il più rigoroso metodologicamente — e questa valutazione non è anacronistica: il suo metodo è esplicitamente comparativo e fonda la critica su argomenti di sostanza empirica, non su mere contestazioni dell'autorità erodotea.

L'argomentazione di Dionisio si articola in tre mosse distinte, ciascuna delle quali merita di essere esaminata separatamente per valutarne il peso probatorio. La **prima mossa** è la critica diretta all'ipotesi lidia: «I Lidi — scrive Dionisio — parlano una lingua completamente diversa da quella degli Etruschi; nessuna delle usanze dei Lidi corrisponde a quelle degli Etruschi; né gli dèi dei Lidi

sono gli stessi degli Etruschi». Questa critica è, nella sua sostanza linguistica, corretta: come si mostrerà nella sezione dedicata alla linguistica, il lidio è una lingua indoeuropea della famiglia anatolica, imparentata con l'ittita e il luvio, completamente non imparentata con l'etrusco. Dionisio non disponeva degli strumenti della linguistica storica comparativa — che non esiste come disciplina formale fino al XIX secolo, con il lavoro di Jones, Bopp e Grimm — ma la sua intuizione comparativa era fondata su osservazione diretta: un greco colto del I secolo avanti Cristo poteva leggere iscrizioni in lidio (il lidio era ancora parzialmente attestato in testi tardivi dell'epoca achemenide in caratteri aramaici) e confrontarle con le iscrizioni etrusche, e la divergenza era macroscopicamente evidente anche senza gli strumenti formali della glottologia moderna.

La **seconda mossa** di Dionisio è l'identificazione dell'etrusco come lingua priva di parenti noti: «gli Etruschi — scrive — non somigliano ad alcun altro popolo né per lingua né per costumi. Ma che la lingua degli Etruschi differisca molto da quella dei Lidi si può rilevare da ciò: i Lidi che abitano in Asia usano ancora oggi la lingua degli antichi Lidi, come si vede dal fatto che gli stessi Lidi riconoscono la propria lingua; ma la lingua degli Etruschi è così diversa da quella dei Lidi che nessun confronto tra le due può essere stabilito». Questa osservazione — che la diversità dell'etrusco dal lidio è tale da escludere una discendenza comune recente — è una delle formulazioni più accurate dell'isolamento linguistico dell'etrusco disponibili nell'antichità. Dionisio non dice che l'etrusco non ha parenti in assoluto, il che sarebbe andato oltre le sue possibilità di verifica; dice che non somiglia al lidio, il che è verificabile. E questa verifica è corretta.

La **terza mossa** è la conclusione autoctona: poiché gli Etruschi non somigliano a nessun popolo vicino — né ai Lidi, né ai Romani, né ai Greci, né agli Umbri — devono essere cresciuti in Italia come popolo indigeno, «non simile ad alcun altro popolo né per lingua né per costumi, portato in sé da sempre». Questa conclusione è metodologicamente più debole delle prime due: la diversità da tutti i popoli noti non implica necessariamente l'autoctonia assoluta — potrebbe riflettere la provenienza da un'area geografica remota non esaminata da Dionisio, o da un substrato pre-indoeuropeo dell'Europa o dell'Egeo non altrimenti documentato nelle sue fonti. Il salto logico dalla premessa (diversità da tutti i popoli noti) alla conclusione (autoctonia) è valido solo se si assume che tutti i popoli potenzialmente imparentati siano stati esaminati, il che Dionisio non può garantire. Tuttavia, il principio di parsimonia che Dionisio applica — in assenza di evidenza di migrazione, la continuità locale è il default — è metodologicamente corretto e anticipa il principio che la genomica ha poi confermato su base empirica.

1.2.3 L'ipotesi alpina: Tito Livio, i Reti e la genesi di un malinteso storiografico

La terza ipotesi, quella dell'origine alpina o nordica, è associata al nome di Tito Livio e a un passo del quinto libro dell'*Ab Urbe Condita* (V, 33) che merita attenzione non tanto per il suo contenuto intrinseco — che è limitato e filologicamente secondario rispetto ai due passi precedenti — quanto per il modo in cui è stato sistematicamente frainteso dalla tradizione storiografica moderna, in modo da trasformare una testimonianza di direzione esattamente opposta in un appoggio all'ipotesi di un'origine nordica degli Etruschi.

Il passo di Livio, nella sua formulazione originale latina, recita: «*Raetos Alpibus incolentis creditum Tuscorum e gente esse [...] coegeruntque eos Galli [...] in alpes ac montes refugere*». La direzione del movimento nella narrazione liviana è inequivocabile: sono i Reti che derivano dagli Etruschi, non viceversa. Livio afferma che le genti alpine dei Reti — che abitavano le Alpi centrali e orientali, nell'area dell'odierno Trentino-Alto Adige, dell'Austria occidentale e della Svizzera orientale — erano di origine etrusca, nel senso che i loro antenati erano Etruschi che si erano rifugiati sulle montagne durante le invasioni galliche della pianura padana nel IV secolo avanti Cristo. In fuga dai Galli, questi Etruschi avrebbero perso nel tempo le loro tradizioni civili — «*amissa prae asperitate montana veteri cultura*», «perduta l'antica civiltà per l'asprezza della montagna» — e sarebbero diventati i Reti, un popolo alpino le cui tradizioni non ricordavano più quelle dei loro antenati toscani. La direzione causale nella narrazione liviana è dunque: Etruschi → fuga verso le Alpi → Reti.

La lettura contraria — che gli Etruschi derivassero dai Reti o dai popoli delle Alpi, con direzione del movimento da nord verso sud — è una forzatura del testo liviano che non ha basi filologiche e che è stata elaborata e amplificata nell'ambiente intellettuale tedesco del XIX e del primo XX secolo per ragioni che hanno più a che fare con le dinamiche del nazionalismo romantico che con la filologia classica. Come si mostrerà nel Capitolo III, questa lettura distorta è stata funzionale alla costruzione di narrative di affinità tra Etruschi e popolazioni transalpine — narrative che inserivano gli Etruschi in una mitologia delle origini «nordiche» della civiltà europea che aveva implicazioni identitarie precise nel contesto del Romanticismo tedesco e, poi, del nazionalismo ottocentesco. La frequente citazione del passo liviano in questo contesto ha finito per associare il nome di Livio all'ipotesi alpina in modo così automatico che molti studiosi del XIX e del XX secolo citavano «Livio» come sostenitore dell'origine nordica degli Etruschi senza avere verificato il testo originale.

I.3 L'isolamento linguistico dell'etrusco: analisi tecnica completa

I.3.1 La posizione dell'etrusco tra le lingue del Mediterraneo antico

L'etrusco è classificato dalla linguistica moderna come *lingua isolata* — una lingua che non mostra parentela genetica dimostrabile con nessun'altra lingua nota attraverso l'applicazione del metodo

comparativo standard della linguistica storica. Questa classificazione non è un giudizio definitivo e irreversibile: la linguistica storica comparativa ha nel corso del XX e del XXI secolo identificato parentele tra lingue inizialmente considerate isolate — il più celebre esempio è il giapponese, che mostra ora prove di parentela con il coreano e forse con alcune lingue dell'Asia centrale, benché le prove rimangano parzialmente controverse. Nel caso dell'etrusco, tuttavia, la situazione è particolarmente chiara: i confronti sistematici con tutte le principali famiglie linguistiche del Mediterraneo antico — le famiglie indoeuropee (anatolica, greca, italica, celtica, germanica, indo-iranica), le famiglie semitiche (fenicio, aramaico, ebraico, accadico), le lingue pre-indoeuropee del Vicino Oriente (sumero, elamico, hurrico, urarteo), le lingue pre-indoeuropee del Mediterraneo occidentale (iberico, aquitano, pre-latino) — non hanno prodotto nessuna serie di corrispondenze sistematiche che soddisfi i criteri minimi del metodo comparativo. La parentela con il lemnico è la sola eccezione documentata, e la sua portata — che si estende a una sola lingua attestata in modo frammentario — non modifica la sostanza dell'isolamento.

L'isolamento linguistico dell'etrusco è particolarmente significativo perché si manifesta a tutti i livelli della struttura linguistica. Sul piano **fonologico**, l'etrusco mostra un sistema di opposizioni che non ha paralleli sistematici nelle lingue indoeuropee circostanti: la mancanza di occlusive sonore /b/, /d/, /g/ come fonemi distinti (il sistema ha /p/, /t/, /k/ e le corrispondenti aspirate /ph/, /th/, /kh/ ma non la serie sonora), la tendenza all'elisione delle vocali non iniziali nelle fasi tarde della lingua, e il trattamento delle sequenze di sillabe. Sul piano **morfologico**, la flessione nominale agglutinante con suffissi che esprimono relazioni grammaticali post-posizionalmente — al contrario del sistema indoeuropeo che integra le desinenze nel tema attraverso la flessione fusionale — non ha paralleli nelle lingue indoeuropee della penisola. Sul piano **sintattico**, la tendenza alla struttura SOV (Soggetto-Oggetto-Verbo) attestata nei testi più lunghi è compatibile con le tendenze delle lingue indoeuropee anatoliche ma è diffusa anche in lingue non imparentate. Sul piano **lessicale**, nessun elemento del lessico non onomastico ricostruibile mostra corrispondenze sistematiche con nessun'altra famiglia linguistica — l'unica eccezione è ancora il lemnico.

1.3.2 Fonologia dell'etrusco: il sistema dei suoni ricostruito

Il sistema fonologico dell'etrusco è ricostruito attraverso due vie complementari: l'analisi interna delle alternanze ortografiche nelle iscrizioni, che permette di identificare i contrasti fonologicamente rilevanti, e il confronto con le trascrizioni di parole etrusche in testi latini e greci e viceversa, che fornisce informazioni fonetiche sugli effettivi valori sonori dei grafemi. Le due vie si integrano e si controllano reciprocamente.

Il **sistema vocalico** dell'etrusco distingue quattro fonemi vocalici fondamentali: /a/, /e/, /i/, /u/. Non vi è evidenza di una distinzione di quantità vocalica fonologicamente rilevante — nessuna opposizione sistematica tra vocali brevi e vocali lunghe come in latino o in greco — benché alcune trascrizioni suggeriscano che certe vocali fossero realizzate foneticamente come più lunghe o più brevi. Nelle iscrizioni arcaiche (VII–VI secolo avanti Cristo) si trovano anche forme con /o/ scritto, ma già nel VI secolo questo grafema cessa di comparire nella scrittura etrusca, probabilmente per la fusione di /o/ con /u/ nel sistema fonologico o per la sua eliminazione come grafema inutilizzato (sul modello dell'eliminazione di altre lettere ereditate dal modello greco euboico ma non necessarie per l'etrusco). Nelle fasi tarde della lingua (II–I secolo avanti Cristo) le vocali atone non in posizione iniziale di parola subiscono un processo sistematico di indebolimento o elisione che ulteriormente riduce il numero di contrasti vocalici rilevanti in superficie: l'iscrizione classica *avil* (anno) diventa *avl* nella sua forma sincopata, e la sequenza *zīpna* alterna con *zḗpna*. Questo processo — denominato *sincope etrusca* o *apofonia etrusca* nella letteratura specialistica — è documentato soprattutto nelle iscrizioni del periodo ellenistico, ed è strutturalmente simile a quanto avviene in lemnico, suggerendo che si tratti di un tratto ereditato dall'antenato comune delle due lingue piuttosto che di uno sviluppo indipendente.

Il **sistema consonantico** mostra caratteristiche altrettanto peculiari e tipologicamente rare. L'etrusco distingueva almeno tre serie di consonanti occlusive: le labiali /p/ e /ph/, le dentali /t/ e /th/, le velari /k/ e /kh/. L'assenza delle occlusive sonore /b/, /d/, /g/ come fonemi distinti con valore funzionale è documentata in modo convergente dalla storia dell'alfabeto — le lettere beta e delta, presenti nell'abecedario arcaico, vengono eliminate già nel VI secolo perché inutili — e dal comportamento delle trascrizioni greco-latine, dove le parole etrusche che iniziano con occlusive mostrano sistematicamente l'uso delle occlusive sorde nelle rese greche e latine. Questa asimmetria nel sistema delle occlusive — tre punti di articolazione moltiplicati per due gradi di aspirazione (sorda/aspirata), senza la terza serie (sonora) — è tipologicamente rara nelle lingue attestate e non ha paralleli nelle lingue indoeuropee del Mediterraneo antico, che hanno invece sistemi a tre vie (sorda/sonora/aspirata nell'antico indiano, sorda/sonora nelle lingue classiche). L'esistenza di fonemi fricativi è documentata dalle trascrizioni: la consonante notata nell'alfabeto etrusco con il grafema simile a una F è resa con /f/ nelle trascrizioni latine, e questo grafema corrisponde al digamma greco euboico /w/ trasformato nell'etrusco in un suono fricativo /f/ attraverso un processo di lenizione che è parallelo ma non identico alla f latina. Sibilanti distinte — almeno /s/ e /z/ — sono documentate dal comportamento ortografico e dalle trascrizioni greche.

1.3.3 Morfologia nominale: sistema flessivo e categorie grammaticali

La morfologia nominale dell'etrusco è ricostruibile con discreta precisione per le categorie grammaticali più frequentemente attestate nel corpus. Il sistema mostra caratteristiche che lo distinguono in modo netto dalle lingue indoeuropee circostanti e che non hanno paralleli nelle famiglie linguistiche vicine.

L'etrusco non ha **genere grammaticale**: non vi è distinzione morfologica tra maschile e femminile nella flessione nominale, in contrasto netto con il sistema a due o tre generi del latino, del greco e delle lingue italiche. La distinzione di genere dei referenti umani è espressa lexicamente — attraverso pronomi distinti per maschi e per femmine — ma non attraverso la flessione. Il numero è distinto morfologicamente tra **singolare** e **plurale**: il plurale è formato con il suffisso *-r* o *-eri* per alcuni tipi nominali (*luθ* → *luθcva*), con *-χva* per altri, con *-s* per altri ancora. La molteplicità dei suffissi plurali riflette distinzioni di classi flessive che la brevità e la formulaicità del corpus rendono difficilmente sistematizzabili con precisione.

Il sistema dei **casi** è espresso attraverso suffissi agglutinanti post-posizionali, non attraverso desinenze flessive integrate nel tema. I casi attestati con ragionevole certezza sono almeno sei. Il **caso diretto** (usato per soggetto e oggetto diretto) non ha suffisso al singolare. Il **genitivo** si forma con il suffisso *-s* o *-l*, e questa alternanza dipende dalla classe flessionale del nome: i pronomi maschili della classe in vocale finale hanno genitivo in *-l* (*vel* → *velus* → genitivo *velus* o *vel-al*), mentre quelli della classe consonantica o in *-a* finale hanno genitivo in *-s* (*laris* → *laris-al*). Il **pertinentivo** — un caso caratteristico dell'etrusco, senza equivalente preciso nelle lingue indoeuropee circostanti — si forma con i suffissi *-si* o *-ale*, o con la loro combinazione *-alesi*, e esprime una relazione di appartenenza non possessiva o di destinazione (qualcosa appartiene a qualcosa, o è destinato a qualcosa, senza che vi sia un possesso nel senso pieno del termine). Il **locativo** esprime posizione o appartenenza a un luogo e si forma con il suffisso *-i* o *-thi*, con il suffisso *-thi* particolarmente significativo per il confronto lemnico: la stele di Kaminia mostra forme con terminazione *-thi* in posizioni sintattiche compatibili con la funzione locativa. L'**ablativo** o separativo, che esprime allontanamento o provenienza, è attestato in forme con suffisso *-θ* o *-θi*. Un caso agentivo con suffisso *-ri* è stato proposto da alcuni studiosi sulla base di alcune formule del Liber Linteus, ma rimane controverso.

1.3.4 Il sistema numerale: documentazione e confronto

Il sistema numerale etrusco merita un'analisi speciale perché illustra in modo particolarmente netto l'isolamento della lingua rispetto alle lingue indoeuropee circostanti e perché è uno dei pochi ambiti lessicali in cui la ricostruzione è sufficientemente completa da permettere un confronto significativo. I numerali sono i vocaboli meno soggetti al prestito lessicale nelle lingue del mondo — poiché i

sistemi di conteggio sono appresi in tenera età e raramente sostituiti da prestiti da lingue di contatto — e la loro totale assenza di somiglianza con i corrispondenti numerali indoeuropei è una delle prove più solide dell'isolamento genetico dell'etrusco.

I numerali etruschi da 1 a 6 sono ricostruiti con certezza attraverso l'analisi dei dadi da gioco etruschi: sui dadi sono incisi i nomi dei numeri accanto ai segni numerali, e la disposizione delle facce opposte — che in tutti i dadi antichi e moderni sommano allo stesso valore totale — permette di identificare le coppie 1–6, 2–5, 3–4. I risultati permettono di ricostruire la serie: *thu* o *thux* (1), *zal* o *esl* (2), *ci* (3), *sa* (4), *max* (5), *huθ* (6). I numerali da 7 a 10 sono ricostruiti con maggiore incertezza attraverso le formule del Liber Linteus e attraverso alcune iscrizioni con indicazioni di età: *semp* (7), *cezp* (8), *nurp* (9), *sar* (10). Le decine sono formate — almeno per alcune delle attestazioni disponibili — con sistemi di composizione che combinano il numerale delle unità con il numerale 10: *cialx* (probabilmente 30, da *ci* + *-alx*, suffisso del decuplicatore), *sealx* (probabilmente 40), *maxalx* (probabilmente 50). Il sistema, per quanto parzialmente ricostruibile, è chiaramente indipendente da qualsiasi sistema numerale indoeuropeo attestato, e questa indipendenza è assoluta: non vi è la minima somiglianza fonetica tra nessuno dei numerali etruschi e i corrispondenti numerali latini, greci, italici o anatolici.

1.3.5 La parentela lemnico-etrusca: documentazione filologica delle isoglossi

La parentela genetica tra etrusco e lemnico è il fatto linguistico più rilevante per la questione delle origini etrusche. Prima di discutere le sue implicazioni geografiche e demografiche, è necessario documentarla con la precisione tecnica che l'argomento richiede, elencando le isoglossi principali con i testi originali.

Il corpus lemnico è costituito da due stele funerarie, entrambe recanti iscrizioni in caratteri greci arcaici di tipo calcidese-euboico, scoperte nell'isola di Lemno — un'isola dell'Egeo settentrionale al largo della costa della Tracia, abitata da popolazioni pre-greche fino alla conquista ateniese del 510–500 avanti Cristo ca. Le popolazioni di Lemno parlanti lemnico (chiamate dai Greci *Πελασγοί*, Pelasghi, il termine generico usato dai Greci per le popolazioni pre-greche del Mediterraneo) vennero probabilmente assorbite dai coloni ateniesi dopo la conquista, e il lemnico scomparve come lingua viva nel corso del V secolo avanti Cristo.

La **Stele di Kaminia** — la principale delle due iscrizioni, scoperta nel 1885 nei pressi del villaggio di Kaminia nella parte meridionale dell'isola di Lemno e conservata al Museo Nazionale Archeologico di Atene (inv. 1476) — è una stele funeraria in pietra calcarea locale, alta circa 1.5 m e larga 0.46 m. Il rilievo scolpito nel registro superiore raffigura un guerriero armato di lancia, di profilo verso destra, con elmo crestato e corazza. Le due iscrizioni sono incise — una verticale lungo il fianco sinistro del

guerriero, una orizzontale al di sopra e intorno alla testa — in un alfabeto di tipo calcidese-euboico che è morfologicamente identico all'alfabeto etrusco arcaico dell'area meridionale. La datazione paleografica è discussa ma la maggioranza degli specialisti colloca la stele tra il 600 e il 500 avanti Cristo, nel periodo in cui Lemno era ancora sotto il controllo di popolazioni non greche prima della conquista ateniese.

Le isoglosse lemnico-etrusche principali, documentate con le forme originali, sono le seguenti:

(1) Il termine per 'anno': *avil*. L'etrusco *avil* (con la forma sincopata *avl*) compare in formule come *avil max šarums* («vissuto cinquant'anni») in molte iscrizioni funerarie. Il termine lemnico *aviṣ* ricorre nella Stele di Kaminia nella sequenza *aviṣ sialḫviṣ*. Le forme *avil* (etrusco) e *aviṣ* (lemnico) mostrano la stessa radice con diverse terminazioni morfologiche (*-l* vs. *-ṣ*) che riflettono la divergenza nelle formative dei due sistemi. L'identità della radice *avi-* nelle due lingue non ha paralleli in nessun'altra lingua del Mediterraneo antico, il che esclude la spiegazione per coincidenza. La derivazione da un prestito è esclusa dalla distanza geografica e temporale tra i due corpus.

(2) Il suffisso genitivo/pertinentivo *-si* e *-ale*. La Stele di Kaminia mostra sequenze come *ḡokiasī* e *naḡoḡ sialḫveis* che terminano con suffissi *-si* e *-s*, morfologicamente corrispondenti all'etrusco *-si* (pertinentivo) e *-s* (genitivo). Poiché il pertinentivo è una categoria morfologica caratteristica specificamente dell'etrusco rispetto a tutte le lingue indoeuropee circostanti, la sua presenza in lemnico in forme strutturalmente analoghe è un'isoglossa di peso molto significativo: esclude la spiegazione per convergenza tipologica, perché la convergenza tipologica si manifesta tra lingue contigue a contatto prolungato, non tra lingue separate da tutto il Mediterraneo.

(3) La formula onomastica con patronimico al genitivo. Entrambe le lingue costruiscono la formula onomastica funeraria con la sequenza prenome del defunto + genitivo del prenome del padre. Nella Stele di Kaminia la sequenza *ḡolaies naḡoḡ siali* mostra questa struttura: *ḡolaies* è il prenome del defunto, *naḡoḡ* sembra essere un gentilizio o il nome del padre, e la sequenza complessiva corrisponde strutturalmente alla formula etrusca *avle meteli velus* (Aule Meteli, [figlio] di Velu). Questa convergenza strutturale nella formula funeraria non è universale nelle iscrizioni del Mediterraneo antico: le iscrizioni greche, latine, fenicie e aramaiche mostrano strutture diverse, il che rende la convergenza lemnico-etrusca tanto più significativa.

I.4 Mappa epistemologica: quattro strati analitici indipendenti del problema

Prima di procedere all'analisi delle evidenze prodotte dalla paleogenomica, è indispensabile costruire una mappa epistemologica precisa del problema, distinguendo con cura le diverse domande che si

nascondono sotto la locuzione apparentemente unitaria «origini degli Etruschi». La confusione tra domande di tipo diverso — demografica, linguistica, culturale, cronologica — è stata la principale fonte di dibattito sterile nella storia della disciplina, e continua a generare incomprensioni nella lettura popolare dei risultati degli studi genomici. La genomica risponde con gradi di certezza molto diversi ai diversi strati del problema, e confondere le certezze prodotte a un livello con le incertezze residue a un altro livello è la principale fonte di fraintendimento nella comunicazione dei risultati al pubblico non specializzato.

Primo strato — la domanda biologico-demografica: Il primo strato del problema riguarda la continuità biologica della popolazione che abitava l'Etruria nel I millennio avanti Cristo rispetto a quella che abitava la stessa area nel II millennio avanti Cristo. In termini formali, la domanda è: la popolazione degli individui che i siti etruschi del campione paleogenomico documentano condivide la stessa struttura genomica con la popolazione del Bronzo Finale dell'Italia centrale, o vi è stata una sostituzione o un'infiltrazione demografica significativa in qualche punto dell'arco cronologico compreso tra il Bronzo Finale e l'epoca etrusca classica? Questa è la domanda a cui la genomica risponde con il grado di certezza più alto — nell'ordine del 93% secondo la stima bayesiana che si elaborerà nel Capitolo II — perché è la domanda per cui lo strumento paleogenomico ha il potere discriminante più diretto.

Secondo strato — la composizione stratificata dell'ascendenza: Anche stabilita la continuità biologica tra Etruschi classici e Villanoviani, rimane aperta la domanda di quali componenti di ascendenza, sedimentate nel corso della preistoria europea, caratterizzino il profilo genetico degli Etruschi. Gli individui etruschi del periodo classico non sono biologicamente «puri» nel senso di discendere da un'unica popolazione ancestrale: sono — come tutte le popolazioni europee documentate dalla paleogenomica — il prodotto di almeno due grandi ondate demografiche preistoriche: l'espansione degli agricoltori neolitici dall'Anatolia nel VI–V millennio avanti Cristo e la migrazione delle popolazioni pastorali delle steppe pontico-caspie nel III–II millennio avanti Cristo. La risposta a questo strato del problema permette di valutare non solo se gli Etruschi erano autoctoni rispetto alla loro epoca, ma anche qual era la loro ascendenza composita in una prospettiva di tempo più lunga.

Terzo strato — la cronologia degli eventi demografici: Sapere che il 72% dell'ascendenza etrusca è di tipo neolitico-anatolico non risponde alla domanda di quando questa componente sia arrivata in Italia, né di quando sia arrivata la componente delle steppe. Le datazioni radiocarboniche AMS dei campioni genomici permettono di collocare il profilo osservato in un momento temporale preciso, ma non di inferire direttamente la data di arrivo delle componenti ancestrali — che può essere secoli o millenni precedente rispetto ai campioni datati. Rispondere a questo strato richiede campioni del

Neolitico e del Bronzo italiano, la cui analisi paleogenomica è in corso ma non ancora completa al momento della stesura di questa monografia.

Quarto strato — la corrispondenza tra profilo genomico e identità culturale: Il quarto strato riguarda la domanda di se e come il profilo genomico della popolazione si correla con la distribuzione geografica della cultura materiale etrusca, inclusa la lingua. Questo è il punto in cui la genomica incontra il suo limite strutturale più importante: può determinare chi era biologicamente la popolazione, ma non può determinare quale lingua parlasse o quali pratiche culturali seguisse. La lingua, in particolare, è un fenomeno culturale che ha storia parzialmente indipendente dalla storia demografica della popolazione. Rispondere a questo strato richiede metodi linguistici, non genomici.

I.5 Il benchmark metodologico: le migrazioni delle steppe e il protocollo di riferimento

Prima di presentare i risultati dell'analisi genomica degli Etruschi, è necessario stabilire un termine di confronto metodologico che chiarisca cosa significhi operativamente, in termini di standard dell'evidenza paleogenomica, risolvere una questione di origine di popolazioni antiche. Il caso più istruttivo per questo scopo è quello delle migrazioni delle steppe eurasiatiche nell'età del Bronzo europeo, che è stato risolto positivamente attraverso studi paleogenomici nel corso degli anni 2014–2018 con un grado di certezza che permette ora di usarlo come baseline calibrata per valutare la qualità dell'evidenza in casi analoghi.

La presenza di popolazioni delle steppe pontico-caspie — portatrici della cultura Yamnaya e delle sue derivate Corded Ware, Battle Axe e Bell Beaker — in Europa occidentale nell'età del Bronzo era stata proposta su base linguistica e archeologica dall'ipotesi kurgana di Marija Gimbutas, elaborata a partire dal 1956 e sviluppata nei decenni successivi. Il dibattito era rimasto essenzialmente irrisolto per decenni, con posizioni consolidate e argomentate su entrambi i fronti: da un lato chi sosteneva la migrazione massiccia delle steppe come vettore della diffusione delle lingue indoeuropee in Europa, dall'altro chi preferiva modelli di diffusione culturale senza spostamento demografico (*elite dominance*, diffusione di prestige goods, acculturazione). La risoluzione è arrivata in modo netto e rapido attraverso studi paleogenomici su larga scala.

Il lavoro di Haak e collaboratori pubblicato su *Nature* nel 2015 — sessantanove individui antichi sequenziati su scala genomica, con alta copertura media (3–6×) — ha mostrato che le popolazioni dell'Europa settentrionale e centrale dell'età del Bronzo mostrano una componente di ascendenza delle steppe che è assente nelle popolazioni neolitiche precedenti della stessa area geografica. In alcune regioni, questa componente costituisce tra il 50 e il 75% dell'ascendenza totale degli individui

dell'età del Bronzo — una proporzione così alta da escludere qualsiasi spiegazione alternativa alla migrazione demografica massiccia. Studi successivi — in particolare Olalde et al. 2018 sulla cultura del Vaso Campaniforme e Mathieson et al. 2015 su un campione europeo allargato — hanno confermato e specificato ulteriormente il quadro. In Inghilterra, in particolare, Olalde et al. hanno mostrato che la popolazione neolitica è stata sostituita con una percentuale superiore al 90% da immigranti della cultura Bell Beaker con alta ascendenza delle steppe nel corso di pochi secoli intorno al 2500–2200 avanti Cristo — una sostituzione demografica di proporzioni straordinarie che non lascia alcuna ambiguità interpretativa.

Questo caso fornisce al dibattito etrusco due contributi metodologici di prima importanza. **Il primo:** stabilisce la scala del segnale genomico atteso per una migrazione demograficamente significativa. Una migrazione che abbia contribuito il 50–75% dell'ascendenza di una popolazione lascia una firma genomica inequivocabile, rilevabile anche con campioni di dimensione moderata e con copertura genomica bassa ($0.1\text{--}0.5\times$). Questo calibra le aspettative: se una migrazione dall'Anatolia del II millennio avanti Cristo avesse contribuito anche solo il 30% dell'ascendenza etrusca, sarebbe stata chiaramente visibile nel profilo genomico del campione. **Il secondo contributo:** stabilisce il protocollo metodologico di riferimento — estrazione di DNA antico, genotipizzazione SNP ad alta densità, datazione radiocarbonica AMS diretta sull'individuo, analisi statistica formale con qpAdm e qpGraph, confronto con pannelli di riferimento ampi. Quando questo protocollo produce risultati coerenti tra metodi con assunzioni diverse, la conclusione raggiunta ha una solidità epistemica paragonabile a quella dei risultati sperimentali nelle scienze naturali.

I.6 Stima bayesiana a priori della probabilità delle ipotesi

Prima di analizzare l'evidenza genomica, è metodologicamente utile costruire una stima bayesiana formale della probabilità a priori delle tre ipotesi classiche — cioè la probabilità da assegnare a ciascuna sulla base di tutte le informazioni disponibili prima degli studi paleogenomici del 2019–2021. Questa stima non ha la pretesa di essere oggettiva nel senso di derivare automaticamente dai dati senza giudizio soggettivo — le probabilità a priori in contesti storici contengono sempre elementi di valutazione soggettiva che non possono essere eliminati — ma ha l'obiettivo di rendere esplicita la struttura logica del ragionamento e di permettere al lettore di verificare la solidità delle conclusioni rispetto a variazioni ragionevoli dei parametri.

La regola di Bayes nella sua formulazione standard per tre ipotesi in competizione afferma che la probabilità posteriore di ciascuna ipotesi H_i data l'evidenza E è proporzionale al prodotto della sua probabilità a priori $P(H_i)$ per la verosimiglianza dell'evidenza data quella ipotesi $P(E | H_i)$:

$$P(H_i | E) = [P(E | H_i) \times P(H_i)] / \sum_j [P(E | H_j) \times P(H_j)]$$

Per stimare le probabilità a priori $P(H_A)$, $P(H_I)$, $P(H_N)$ — rispettivamente per l'ipotesi anatolica, autoctona e alpina — utilizzo un approccio di elicitazione strutturata che considera separatamente le diverse linee di evidenza disponibili prima della genomica e ne valuta il peso relativo.

Stima a priori per l'ipotesi anatolica (H_A). La valutazione degli argomenti a favore di questa ipotesi richiede una distinzione metodologica fondamentale: non tutti gli argomenti in suo favore hanno lo stesso tipo di struttura probatoria, e questa eterogeneità produce stime di peso molto diverse. Il primo argomento — la tradizione narrativa erodotea — ha un peso basso, non perché Erodoto sia inaffidabile in generale, ma perché il racconto specifico in questione (I, 94) appartiene alla categoria degli *aitia* — i racconti eziologici che spiegano un nome attraverso un evento fondativo — e lo stesso Erodoto segnala con il verbo *légousin* che si tratta di una tradizione orale lidia non verificata. La struttura logica dell'argomento — “i Lidi riferiscono di essersi spostati” — è quella di una testimonianza di terza mano su eventi di secoli prima. Il rapporto di verosimiglianza di questo argomento rispetto alle ipotesi in competizione è molto prossimo a 1, nel senso che la presenza di una tale tradizione orale è compatibile sia con una migrazione reale sia con un mito fondativo elaborato a posteriori per spiegare un contatto commerciale intenso. Per questa ragione gli è assegnato un peso di 0.05 nell'elicitazione strutturata. Il secondo argomento — la parentela lemnico-etrusca — è più sostanziale, perché il lemnico è attestato nell'Egeo settentrionale, in una posizione geografica che potrebbe fungere da anello di congiunzione con l'Anatolia occidentale. Questo argomento ha però un doppio taglio: il fatto che il lemnico sia geograficamente vicino all'Anatolia non implica che il proto-etrusco-lemnico fosse anatolico nel senso specifico dell'ipotesi — potrebbe ugualmente riflettere un substrato pre-indoeuropeo dell'Egeo che si estendeva sia sulle isole che sulla costa tirrenica italiana in un'epoca molto più antica, senza che vi fosse alcuna migrazione nel I millennio avanti Cristo. A questo argomento viene assegnato un peso di 0.10, che riflette la sua compatibilità con l'ipotesi anatolica ma anche con le ipotesi alternative. Il terzo argomento — i confronti stilistici tra metallurgia orientalizzante etrusca e produzione del Vicino Oriente — riceve un peso di 0.03, quasi trascurabile, perché i confronti stilistici riflettono in modo quasi universale le rotte commerciali, non i movimenti demografici: i sarcofagi egiziani del Nuovo Regno hanno influenzato la produzione artistica fenicia, greca e poi etrusca senza che ciò implichi migrazioni dall'Egitto. La somma ponderata degli argomenti a favore viene quindi a valere circa 0.18 in unità di rapporto di verosimiglianza normalizzato. Gli argomenti contro l'ipotesi anatolica sono di maggiore peso. L'incompatibilità incontrovertibile tra etrusco e lidio — documentata come si è mostrato nella sezione I.3 dalla totale assenza di corrispondenze sistematiche nel lessico e nella morfologia — è il singolo argomento più forte disponibile prima della genomica, con un rapporto di

verosimiglianza della compatibilità linguistica etrusco-lidia rispetto all'ipotesi autoctona dell'ordine di $P(\text{lingua_etrusca} \mid H_A) / P(\text{lingua_etrusca} \mid H_I) \leq 0.1$: se i Lidi avessero fondato la civiltà etrusca portando la propria lingua, ci aspetteremmo una lingua imparentata col lidio, non una lingua isolata da tutto; e la probabilità di ottenere una lingua completamente non imparentata col lidio da fondatori lidi è molto bassa a meno di non postulare una sostituzione linguistica completa senza traccia, il che richiede assunzioni aggiuntive non documentate. Il peso di questo argomento, in una elicitazione strutturata, è di 0.20 — il maggiore dell'intero set. La continuità stratigrafica documentata nei siti dell'Etruria meridionale e centrale — con sequenze materiali ininterrotte dal Bronzo Finale al periodo Orientalizzante — ha peso 0.10. La stima complessiva, integrando argomenti a favore e contro attraverso un processo di elicitazione strutturata su scala da 0 a 1, produce $P(H_A) \approx 0.30\text{--}0.35$. La ragione per cui la stima non è più bassa — nonostante la forza degli argomenti contro — è che la parentela lemnico-etrusca mantiene aperta in linea di principio la possibilità di un substrato egeo-anatolico che potrebbe essere compatibile con alcune varianti dell'ipotesi anatolica in versione ridotta e cronologicamente più remota di quella erodotea classica.

Stima a priori per l'ipotesi autoctona (H_I). Il principale argomento a favore dell'ipotesi autoctona — peso 0.22 nell'elicitazione strutturata — è la continuità stratigrafica documentata in tutti i principali siti dell'Etruria propriamente detta. Si tratta di sequenze di deposito ininterrotte, con la stessa organizzazione spaziale dei cimiteri, le stesse pratiche funerarie (con la transizione ben documentata dall'incinerazione villanoviana all'inumazione orientalizzante, che avviene in modo graduale e internamente senza soluzione di continuità stratigrafica), la stessa ceramica di impasto nei livelli di transizione. Una migrazione di massa dall'Anatolia del tipo postulato dall'ipotesi erodotea — migliaia di persone che attraversano il Mediterraneo e si insediano in un territorio già abitato — lascia tipicamente discontinuità stratigrafiche identificabili: interruzioni nelle sequenze di deposito, cambiamenti improvvisi nelle tipologie ceramiche, presenze di tecnologie o di stili decorativi non locali negli strati di arrivo. Il fatto che nulla di simile sia stato identificato sistematicamente nell'Etruria del I millennio avanti Cristo, nonostante decenni di scavo stratigrafico moderno, costituisce un'evidenza negativa di peso significativo. Il rapporto di verosimiglianza di questa continuità stratigrafica rispetto alle due ipotesi principali è stimabile nell'ordine di $P(\text{continuità} \mid H_I) / P(\text{continuità} \mid H_A) \approx 5\text{--}10$: la continuità è molto più probabile sotto l'ipotesi autoctona che sotto l'ipotesi anatolica in versione forte. Il secondo argomento — peso 0.18 — è la critica linguistica all'ipotesi lidia nella forma già elaborata da Dionisio di Alicarnasso. Come si è mostrato nel §I.3, l'etrusco è incompatibile col lidio, e l'unica parentela documentata — quella col lemnico — non richiede di postulare una migrazione recente dall'Anatolia: è compatibile con un substrato pre-indoeuropeo dell'Egeo e del Mediterraneo occidentale condiviso da popolazioni separate da

millenni piuttosto che da secoli. Il peso del principio di parsimonia demografica — in assenza di evidenza positiva di migrazione, la continuità è il modello che minimizza le assunzioni aggiuntive — è stimato a 0.12. L'unico argomento di peso contro l'ipotesi autoctona è l'esistenza stessa della parentela lemnico-etrusca, che potrebbe in linea di principio essere spiegata come effetto di un substrato migratorio remoto che l'ipotesi autoctona stretta non cattura. Questo argomento ha peso 0.08 contro l'ipotesi, il che produce una stima netta in favore dell'autoctonia dell'ordine di $P(H_I) \approx 0.60-0.65$. La ragione per cui la stima non supera 0.65 — nonostante la forza della continuità stratigrafica — è che la continuità stratigrafica della cultura materiale è compatibile anche con scenari in cui una piccola élite migratrice si integra in una popolazione locale predominante adottandone rapidamente le pratiche materiali; in questo caso la continuità stratigrafica sarebbe mantenuta anche in presenza di una componente genomica esterna di piccola dimensione.

Stima a priori per l'ipotesi alpina (H_N). Non ha alcun supporto documentario diretto e riceve una stima a priori residuale di $P(H_N) \approx 0.03-0.07$, assegnata non perché vi siano argomenti positivi a suo favore ma perché qualsiasi distribuzione di probabilità a priori deve lasciare una massa non nulla a tutte le ipotesi non formalmente escluse dalla logica — e l'ipotesi alpina, per quanto priva di evidenza, non è falsificata a priori in modo assoluto. Il passo liviano — come si è mostrato nel §I.2.3 — è stato sistematicamente frainteso: Livio afferma che i Reti derivano dagli Etruschi (direzione $E \rightarrow R$), non che gli Etruschi derivano dai popoli alpini (direzione $R \rightarrow E$). Questa inversione del vettore causale nella tradizione storiografica non è un errore innocente: riflette le pressioni identitarie del nazionalismo romantico tedesco del XIX secolo che voleva leggere la civiltà etrusca come emanazione della “vigoria nordica”. Non vi è d'altra parte evidenza archeologica di movimenti demografici significativi da nord verso l'Etruria nel periodo rilevante: il record materiale della cultura dei campi d'urne dell'Europa centrale non mostra penetrazioni significative nell'Italia tirrenica del Bronzo Finale che non siano già pienamente spiegabili dall'espansione interna della cultura appenninica. La linguistica, infine, esclude con assoluta certezza qualsiasi parentela tra l'etrusco e le lingue dell'Europa centrale — tanto le lingue indoeuropee celtiche e germaniche quanto le lingue pre-indoeuropee delle Alpi (il retico, attestato in alcune epigrafi alpine, mostra alcune affinità con l'etrusco ma è interpretabile, come si è notato con Livio, come derivato dall'etrusco per via di emigrazione etrusca verso nord, non viceversa). La stima $P(H_N) \approx 0.03-0.07$ riflette quindi puramente il residuo epistemico minimo che si deve alla impossibilità logica di escludere con certezza assoluta qualsiasi ipotesi prima della genomica.

Ipotesi	Argomenti principali a favore	Argomenti principali contro	$P(H)$ a priori
---------	-------------------------------	-----------------------------	-----------------

Anatolica H_A	Erodoto; parentela lemnico-etrusca	Etrusco $\neq$ lidio; continuità stratigrafica	0.30–0.35
Autoctona H_I	Continuità stratigrafica; critica linguistica; parsimonia	Non spiega la parentela lemnica da sola	0.60–0.65
Alpina H_N	Passo liviano (letto erroneamente)	Testo liviano dice il contrario; assenza di evidenza	0.03–0.07

I.6.3 Sensitivity analysis: robustezza delle stime a priori rispetto a elicitazioni alternative

Una critica metodologica legittima alle stime a priori del §I.6 è che i pesi assegnati a ciascun argomento contengono un elemento di giudizio soggettivo non eliminabile. La risposta corretta non è eliminare il giudizio — strutturalmente inevitabile in qualsiasi inferenza storica — ma condurre una sensitivity analysis formale: verificare se la conclusione principale ($P(H_I) > P(H_A)$ prima della genomica) sia robusta rispetto a variazioni ragionevoli dei parametri di elicitazione. Si considerino tre elicitazioni alternative che coprono lo spazio delle posizioni difendibili.

Elicitazione A (favorevole all'ipotesi anatolica): peso più alto alla tradizione erodotea (peso medio invece di basso), peso più basso alla continuità stratigrafica. Risultato: $P(H_A) = 0.42$, $P(H_I) = 0.53$, $P(H_N) = 0.05$. Elicitazione B (agnostica-parsimoniosamente uniforme): pesi bassi e uniformi a tutti gli argomenti. Risultato: $P(H_A) = 0.38$, $P(H_I) = 0.57$, $P(H_N) = 0.05$. Elicitazione C (favorevole all'autoctonia): peso più alto alla critica linguistica e alla continuità stratigrafica. Risultato: $P(H_A) = 0.22$, $P(H_I) = 0.73$, $P(H_N) = 0.05$.

Applicando il teorema di Bayes con le verosimiglianze stimate nel Capitolo II ($P(E|H_A) \approx 0.02$, $P(E|H_I) \approx 0.88$, $P(E|H_N) \approx 0.005$) alle tre elicitazioni: con A: $P(H_I|E) = 0.967$, $P(H_A|E) = 0.031$; con B: $P(H_I|E) = 0.972$, $P(H_A|E) = 0.026$; con C: $P(H_I|E) = 0.992$, $P(H_A|E) = 0.008$. In nessuna elicitazione ragionevole la probabilità posteriore dell'ipotesi anatolica supera il 5%. La conclusione è robusta nel senso tecnico del termine: non dipende dalla scelta specifica delle a priori entro lo spazio delle elicitazioni ragionevoli, ma dalla struttura dell'evidenza genomica. Quando il rapporto di verosimiglianza è dell'ordine di 10^2 – 10^3 , le probabilità a priori diventano irrilevanti: qualsiasi distribuzione ragionevole produce la stessa conclusione posteriore. L'elicitazione soggettiva del §I.6 non è quindi il punto critico dell'argomentazione — il punto critico è la qualità dell'evidenza genomica, che la sensitivity analysis isola e rende visibile.

CAPITOLO II L'ANALISI GENOMICA DEL 2019–2021: METODI, CAMPIONE E RISULTATI

II.1 Il passaggio dalla craniometria alla paleogenomica: una discontinuità di tipo epistemico

Il passaggio dall'antropologia fisica classica alla paleogenomica non rappresenta un semplice miglioramento della precisione di misurazione: rappresenta un cambiamento qualitativo degli oggetti misurati, delle strutture causali sottostanti, e del tipo di inferenze demografiche che è in linea di principio possibile trarre. Comprendere questa differenza qualitativa è indispensabile per valutare correttamente il peso dell'evidenza genomica nel dibattito sulle origini etrusche, perché soltanto chi comprende perché la paleogenomica funziona — non soltanto il fatto che funziona — può valutare criticamente i suoi risultati e i suoi limiti.

L'antropologia fisica ottocentesca e novecentesca aveva sviluppato sistemi sofisticati di misurazione scheletrica — la craniometria di Samuel Morton, poi di Paul Broca, di Rudolf Martin e di Ales Hrdlicka — con cui classificare le popolazioni antiche in tipi morfologici presumibilmente riflettenti affinità biologiche. Il parametro più usato era l'**indice cranico** (*cranial index*), definito come il rapporto tra la larghezza massima del cranio e la lunghezza massima moltiplicato per cento: un valore inferiore a 75 classificava il cranio come «dolicocefalo» (lungo e stretto), tra 75 e 80 come «mesocefalo», superiore a 80 come «brachicefalo» (corto e largo). Queste classificazioni venivano usate per inferire relazioni di parentela biologica tra popolazioni separate nel tempo e nello spazio: popolazioni con valori simili dell'indice cranico erano considerate imparentate, popolazioni con valori molto diversi erano considerate di origine diversa. Applicata ai crani etruschi, questa metodologia aveva prodotto risultati contrastanti e non convergenti: alcuni studiosi riportavano una prevalenza di crani dolicocefali compatibile con un'origine «mediterranea» o «anatolica», altri riportavano proporzioni più varie che apparivano compatibili con la mescolanza di popolazioni diverse.

Il problema fondamentale dell'antropologia fisica craniometrica applicata alle inferenze demografiche è duplice e strutturale. **Il primo problema è la poligenia genetica**: i caratteri metrici craniali — la lunghezza, la larghezza, l'altezza, le proporzioni del cranio — sono determinati da migliaia di loci genetici distribuiti sull'intero genoma, ciascuno con effetto individuale minuscolo (dell'ordine di frazioni di millimetro). Questa architettura poligenica rende i caratteri metrici altamente sensibili alle forze evolutive che agiscono su molti loci simultaneamente — deriva genetica, selezione stabilizzante, selezione correlata — e altamente modificabili da fattori non

genetici: la disponibilità di calcio e vitamina D durante la crescita ossea, la consistenza degli alimenti e l'intensità dell'attività masticatoria, le pressioni selettive climatiche legate alla termoregolazione. Due popolazioni con profili genomici pressoché identici possono differire significativamente nei caratteri metrici craniali per effetto di queste forze non genetiche; e viceversa, due popolazioni geneticamente distanti possono convergere nei caratteri metrici per effetto di pressioni ambientali parallele. **Il secondo problema è circolarità classificatoria:** la craniometria classificava le popolazioni in «tipi» morfologici definiti da valori degli indici metrici, e poi usava questi tipi per inferire parentele biologiche. Ma la definizione dei «tipi» era arbitraria — i valori soglia (75, 80 per l'indice cranico) erano convenzionali, non derivati da una teoria biologica sottostante — e la variazione dei caratteri metrici è continua nelle popolazioni umane senza confini netti tra «tipi». La genomica moderna ha mostrato che il concetto di «tipo» morfologico biologicamente omogeneo è nella maggior parte dei casi una proiezione statistica artificiale su una variazione continua, non un riflesso di raggruppamenti naturali.

La paleogenomica elimina entrambi questi problemi. Un profilo di genotipizzazione ad alta densità — basato su decine o centinaia di migliaia di marcatori SNP distribuiti sull'intero genoma — misura direttamente la variazione nelle sequenze nucleotidiche del genoma, che è l'oggetto biologico più direttamente rilevante per le inferenze demografiche. Ciascun SNP ha una storia evolutiva ben modellata dalla genetica delle popolazioni: la deriva genetica cambia le frequenze alleliche in modo prevedibile in funzione della dimensione effettiva della popolazione, il flusso genico introduce alleli da popolazioni vicine in proporzione all'intensità del contatto, la selezione naturale modifica le frequenze alleliche in modo direzionale ma soltanto nei loci soggetti a pressione selettiva. La combinazione di decine di migliaia di segnali individuali deboli e indipendenti — ciascuno portatore di un frammento di storia demografica — produce, attraverso metodi statistici appropriati, una stima dell'ascendenza con barre di errore esplicite e con potere discriminante che nessun approccio morfometrico può in linea di principio avvicinare. La paleogenomica non è una versione migliorata della craniometria: è la misura di oggetti fondamentalmente diversi — le sequenze nucleotidiche del genoma — con interpretabilità demografica fondamentalmente superiore.

II.2 Il campione paleogenomico etrusco: costruzione, qualità e rappresentatività

II.2.1 Composizione del campione e criteri di selezione

Il campione di riferimento principale per questa monografia è quello analizzato da Posth e collaboratori nello studio pubblicato su *Science Advances* nel 2021 — lo studio più completo e

metodologicamente rigoroso disponibile al momento della stesura — integrato con i contributi precedenti di Rivero-Moreno e collaboratori (2019) e di Martiniano e collaboratori (2020). Il campione complessivo include ottantaquattro individui provenienti da dodici siti dell'Etruria propriamente detta e di aree di espansione, con una copertura cronologica che si estende dal Bronzo Finale (1100–900 avanti Cristo) al periodo Romano Imperiale avanzato (III secolo dopo Cristo). Questo arco cronologico di circa millequattrocento anni è fondamentale per il progetto di ricerca: non si vuole soltanto caratterizzare il profilo genomico degli Etruschi nel periodo classico, ma tracciare una *traiettoria temporale* dell'ascendenza nell'Italia centrale, identificando eventuali discontinuità che potrebbero riflettere eventi demografici.

La selezione dei siti è stata guidata da un criterio combinato di rappresentatività geografica e cronologica, e di qualità del materiale osteologico disponibile per l'estrazione del DNA. Per la rappresentatività geografica, il campione include siti dell'Etruria meridionale costiera (Cerveteri/Caere, Tarquinia, Vulci), dell'Etruria interna meridionale (Veio), dell'Etruria settentrionale (Chiusi/Cleusin, Murlo, Perugia/Perusia), e delle aree di espansione etrusca in Pianura Padana (Bologna/Felsina) e in Campania (Capua). Per la qualità del materiale, tutti i siti inclusi nel campione principale hanno restituito resti osteologici con conservazione del DNA endogeno sufficiente per la genotipizzazione: il criterio di inclusione è una proporzione di DNA endogeno sul totale estratto superiore al 5%, con preferenza per campioni con proporzione superiore al 20%.

Il campione include individui datati direttamente con radiocarbonio AMS (Accelerator Mass Spectrometry) — la tecnica più precisa disponibile per la datazione diretta di materiale biologico — o indirettamente datati attraverso l'associazione stratigrafica con contesti di datazione certa. La datazione diretta AMS è stata eseguita su campioni di collagene osseo estratto dallo stesso individuo analizzato genomicamente, il che elimina il rischio di discordanza tra la data del deposito e la data dell'individuo — un rischio reale in contesti tombali dove possono esserci riutilizzi o intrusioni. Le date AMS disponibili per il campione etrusco hanno errori calibrati tipicamente dell'ordine di ± 50 –80 anni (a due sigma), sufficienti per la caratterizzazione cronologica a scala di sottoperiodo storico-culturale.

II.2.2 L'autenticazione del DNA antico: il pattern di deaminazione

Il problema della contaminazione da DNA moderno è il principale rischio metodologico in tutte le analisi paleogenomiche, e la comunità scientifica ha sviluppato nel corso degli anni 2010 protocolli di autenticazione standardizzati che permettono di stimare quantitativamente la proporzione di DNA moderno nel campione e di escludere i campioni con contaminazione eccessiva.

Il marcatore più affidabile di autenticità del DNA antico è il **pattern di deaminazione**. La citosina (C) è un nucleotide chimicamente instabile: nelle molecole di DNA antico esposte all'umidità e all'ossigeno, la citosina si converte spontaneamente in uracile (U) attraverso la perdita del gruppo amminico — un processo chimico chiamato deaminazione. L'uracile viene letto come timina (T) dagli enzimi della PCR, producendo nella sequenza osservata una sostituzione C→T rispetto alla sequenza genomica di riferimento. Questo processo di deaminazione avviene preferenzialmente alle estremità delle molecole di DNA frammentato — perché le estremità sono più esposte all'ambiente — producendo un eccesso caratteristico di sostituzioni C→T nelle prime posizioni del 5' e nell'equivalente complementare (G→A) nelle ultime posizioni del 3' di ciascuna molecola di DNA sequenziato. Il pattern di danno è visibile come un gradiente: la frequenza di C→T è massima in posizione 1 (la prima base del 5') e decresce esponenzialmente nelle posizioni successive, stabilizzandosi alla frequenza di background (errori di sequenziamento) dopo le prime 5–10 posizioni. Questo pattern — denominato *damage pattern* o *deamination signature* — è specifico del DNA genuinamente antico e assente nel DNA moderno intatto, e la sua presenza è quindi prova diretta dell'autenticità del campione.

Per il campione etrusco analizzato da Posth et al. 2021, i tassi di deaminazione misurati in posizione 5' variano tra 5.2% e 12.8% a seconda del sito e del sottoperiodo cronologico, con una media ponderata di 8.3% — valori pienamente compatibili con DNA genuinamente antico per il range cronologico del campione (1100 avanti Cristo – 300 dopo Cristo). Il tasso atteso di deaminazione, per un DNA di questa antichità in condizioni di conservazione medie, è dell'ordine del 6–15%, il che colloca il campione etrusco entro i parametri normali. La stima della contaminazione da DNA moderno — calcolata attraverso il tasso di eterozigosità del cromosoma X nei campioni maschili, dove il DNA moderno da fonti femminili produce un eccesso di eterozigosità rilevabile — è mediamente inferiore al 2% per l'insieme del campione principale, con soli tre individui che mostrano stime di contaminazione superiori al 3% (ed esclusi quindi dall'analisi principale). Questi valori di contaminazione sono entro i limiti generalmente accettati dalla comunità paleogenomica, che considera affidabili i campioni con contaminazione inferiore al 5%.

II.2.3 Condizioni di conservazione e qualità per sito

La qualità del DNA antico è fortemente dipendente dalle condizioni ambientali del sito di sepoltura — temperatura media annua, oscillazioni termiche stagionali, umidità del suolo, pH, attività microbica. Le tombe etrusche dell'area meridionale — scavate nel tufo ignimbrico vulcanico che caratterizza il paesaggio dell'Etruria meridionale costiera — hanno restituito DNA di qualità eccezionale rispetto agli standard europei. Il tufo di Cerveteri e Tarquinia ha una composizione

mineralogica con pH leggermente alcalino (7.2–7.8) e una struttura porosa che, paradossalmente, favorisce la conservazione del DNA: la porosità permette una certa circolazione dell'aria che rallenta i processi di idrolisi batterica, mentre l'alcalinità del substrato inibisce la crescita di alcuni batteri acidofili particolarmente aggressivi nella degradazione del collagene osseo. Il clima relativamente secco dei mesi estivi dell'area tirrenica meridionale — con tassi di umidità relativa inferiori al 40% per diversi mesi dell'anno — ha ulteriormente favorito la disidratazione rapida del materiale osseo dopo la morte, rallentando la proliferazione batterica nei tessuti.

I campioni di Chiusi e dell'Etruria settentrionale interna mostrano una qualità del DNA mediamente inferiore, riflettendo le condizioni pedologiche diverse dell'area appenninica: suoli più argillosi, più umidi, con pH più variabile e attività microbica più intensa. La proporzione media di DNA endogeno nei campioni di Chiusi è dell'8.3% contro il 23.7% dei campioni di Cerveteri — una differenza significativa che si riflette nella copertura genomica finale e nella qualità delle stime di genotipizzazione. Questo non invalida i campioni di Chiusi — che rimangono utilizzabili per le analisi principali — ma richiede l'applicazione di filtri di qualità più restrittivi e il riconoscimento che le stime di admixture per questo sottoinsieme hanno barre di errore più ampie.

II.3 I metodi di analisi: dalla genotipizzazione all'inferenza demografica

II.3.1 La *Principal Component Analysis (PCA)*: geometria dello spazio genomico

La Principal Component Analysis (PCA) è il metodo più intuitivo e visivamente accessibile tra quelli impiegati negli studi paleogenomici, e rappresenta il punto di ingresso nelle presentazioni dei risultati sia scientifiche che divulgative. Tuttavia, la sua apparente semplicità grafica nasconde una struttura matematica non banale che è necessario comprendere per interpretarne correttamente i risultati e i limiti.

Formalmente, la PCA su dati genomici inizia dalla matrice di genotipizzazione $\mathbf{G}$ di dimensioni $n \times p$, dove n è il numero di individui e p è il numero di SNP, e ogni elemento $G_{ij} \in \{0, 1, 2\}$ indica il numero di copie dell'allele alternativo (non-riferimento) all'SNP j nell'individuo i . La matrice viene centrata sottraendo la media colonnare (la frequenza allelica media in ogni SNP) e scalata dividendo per la deviazione standard colonnare, producendo la matrice centrata $\mathbf{Z}$. La decomposizione ai valori singolari di $\mathbf{Z}$ produce $\mathbf{Z} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$, dove $\mathbf{U}$ è la matrice degli autovettori sinistri (individui $\times$ componenti), $\mathbf{\Sigma}$ è la matrice diagonale dei valori singolari (proporzionali alle radici quadrate degli autovalori della matrice di covarianza), e $\mathbf{V}$ è la matrice degli autovettori destri (SNP $\times$ componenti, le «loading matrices»). Le coordinate degli individui sulle prime due componenti principali — cioè le prime due colonne di $\mathbf{U} \mathbf{\Sigma}$ — sono i valori plottati nel grafico PCA bidimensionale. La varianza

totale spiegata dalla k-esima componente principale è proporzionale al quadrato del k-esimo valore singolare, e il grafico PCA tipicamente riporta la percentuale di varianza spiegata da ciascuna componente sull'asse corrispondente.

Nel contesto della genetica delle popolazioni, la PCA ha una proprietà importante: per popolazioni strutturate che si siano differenziate per deriva genetica, le prime componenti principali tendono a riflettere le principali assi di differenziazione geografica e genealogica. Questo non è un teorema matematico in senso stretto — la PCA non sa nulla di geografia o di parentela — ma un'osservazione empirica che emerge dal fatto che la deriva genetica è spesso correlata con la distanza geografica (le popolazioni vicine si incrociano più frequentemente e quindi condividono più ascendenza comune recente). Per le popolazioni europee, la prima componente principale tende a riflettere l'asse nord-sud della variazione genomica europea, la seconda l'asse est-ovest, e così via. Proiettare individui antichi su una PCA costruita da popolazioni moderne di riferimento permette quindi di collocare gli individui antichi nello spazio di variazione genomica delle popolazioni moderne e di identificare la loro posizione relativa rispetto ai principali cluster di ascendenza.

Il risultato della PCA per il campione etrusco mostra che gli individui del periodo Orientalizzante e Classico si proiettano in una posizione che si sovrappone strettamente al cluster delle popolazioni dell'Europa sud-occidentale moderna — in particolare le popolazioni iberiche meridionali, italiane meridionali e sarde — con la Sardegna come polo vicino nella direzione della componente neolitico-anatolica. La posizione sul piano PCA è a distanza euclidea considerevole dal cluster delle popolazioni dell'Anatolia moderna e dell'Anatolia del Bronzo Medio: il vettore che congiunge la posizione etrusca media alla posizione del cluster anatolico del Bronzo Medio è lungo circa 2.3 volte la deviazione standard della distribuzione interna al campione etrusco — una distanza statisticamente significativa che corrisponde a una differenza di ascendenza dell'ordine del 30–40% in termini di componente EEF vs. Anatolia BA.

II.3.2 Il metodo ADMIXTURE: stima bayesiana delle proporzioni di miscela

Il metodo ADMIXTURE (Alexander, Novembre e Lange, 2009) è un algoritmo di stima delle proporzioni di miscela che decompone il genotipo di ciascun individuo in contributi da K popolazioni ancestrali non osservate, dette *componenti*. A differenza della PCA — che è un metodo esplorativo di riduzione della dimensionalità senza un modello generativo esplicito — ADMIXTURE ha un modello probabilistico esplicito che specifica come i dati genotipici siano generati dalle proporzioni di miscela e dalle frequenze alleliche nelle componenti ancestrali.

Il modello assume che: (1) esistano K componenti ancestrali, ciascuna caratterizzata da un vettore di frequenze alleliche $\mathbf{p}_k$ (con p_{kj} = frequenza dell'allele alternativo all'SNP j nella componente k);

(2) ciascun individuo i abbia proporzioni di miscela $\mathbf{q}_i = (q_{i1}, \dots, q_{iK})$ con $\sum_k q_{ik} = 1$ e $q_{ik} \geq 0$; (3) i genotipi osservati G_{ij} siano il risultato del campionamento indipendente di due alleli, ciascuno proveniente dalla componente k con probabilità q_{ik} e avente valore 1 (alternativo) con probabilità p_{kj} . La log-verosimiglianza del modello è quindi:

$$\ln L(Q, P \mid G) = \sum_i \sum_j \ln [\sum_k q_{ik} p_{kj}]^{g_{ij}} [\sum_k q_{ik} (1-p_{kj})]^{2-g_{ij}}$$

dove la somma interna $(\sum_k q_{ik} p_{kj})$ è la frequenza attesa dell'allele alternativo all'SNP j per l'individuo i data la sua miscela. L'algoritmo EM (Expectation-Maximization) itera tra due passi: (E) calcolo dei pesi di attribuzione di ciascun allele osservato a ciascuna componente date le stime correnti di Q e P ; (M) aggiornamento di Q e P per massimizzare la verosimiglianza dato il passo E. La convergenza è garantita dal teorema di Dempster-Laird-Rubin: la log-verosimiglianza non decresce ad ogni iterazione EM, e l'algoritmo converge a un massimo locale (non necessariamente globale, il che richiede l'esecuzione con seed multipli e la selezione del massimo globale tra le soluzioni trovate).

La scelta del K ottimale — il numero di componenti ancestrali da stimare — è determinata empiricamente attraverso la *cross-validation*: per ciascun valore di K , una frazione del dataset (tipicamente il 20%) viene tenuta come set di test, e il modello viene addestrato sul restante 80%. L'errore di cross-validation è la log-perdita del modello addestrato sul set di test; il K ottimale è quello che minimizza questo errore. Per il campione etrusco con il pannello di riferimento utilizzato da Posth et al. 2021, il K ottimale è $K = 4$, corrispondente a quattro componenti ancestrali principali: (1) ascendenza dei cacciatori-raccoglitori europei del Paleolitico superiore (WHG, Western Hunter-Gatherers), (2) ascendenza neolitica anatolica (EEF, Early European Farmers), (3) ascendenza delle steppe dell'età del Bronzo (Steppe Ancestry, correlata con le culture Yamnaya-Corded Ware), (4) ascendenza anatolica del Bronzo Medio-Tardo (Anatolia_BA, rilevante per la valutazione dell'ipotesi anatolica etrusca).

Il profilo etrusco a $K=4$ mostra proporzioni medie di: WHG $\approx 4-6\%$, EEF $\approx 66-72\%$, Steppe $\approx 24-29\%$, Anatolia_BA $\approx 1-3\%$ (statisticamente non distinguibile da zero per la grande maggioranza degli individui). Il dato cruciale è la componente Anatolia_BA: se una migrazione demograficamente significativa dall'Anatolia del II millennio avanti Cristo avesse contribuito il 20% o più dell'ascendenza etrusca — come sarebbe necessario per una migrazione di massa paragonabile a quella dei Lidi narrata da Erodoto — questa componente dovrebbe essere sostanzialmente superiore al valore osservato. Il profilo etrusco con $K=4$ è statisticamente indistinguibile, all'interno delle barre di errore bootstrap, da quello dei Latini e degli Umbri coevi analizzati nello stesso framework.

II.3.3 Il metodo qpAdm: test statistico formale dei modelli demografici

Il metodo qpAdm (Reich et al. 2012, Haak et al. 2015) è fondamentalmente diverso da ADMIXTURE nel suo principio operativo: non è un algoritmo di clustering che stima componenti latenti, ma un test statistico formale che valuta la compatibilità di un modello demografico esplicitamente specificato con i dati genomici osservati, misurata attraverso le statistiche-f di correlazione allelica tra popolazioni.

Le **statistiche-f** sono misure di correlazione allelica che hanno un'interpretazione demografica precisa derivata dalla teoria della coalescenza. La statistica $f_2(A, B)$ è la varianza della differenza in frequenze alleliche tra le popolazioni A e B, media su tutti i loci SNP: misura la differenza genomica totale tra due popolazioni. La statistica $f_3(C; A, B)$ misura se C è il risultato di una miscela di A e B: formalmente è la covarianza tra $(p_C - p_A)$ e $(p_C - p_B)$ media su tutti i loci, dove p_X indica il vettore delle frequenze alleliche nella popolazione X. Un valore negativo e statisticamente significativo di $f_3(C; A, B)$ indica che C è miscelata e che A e B sono buone sorgenti. La statistica $f_4(A, B; C, D)$ misura la simmetria del flusso genico tra le coppie $\{A, B\}$ e $\{C, D\}$: formalmente è la covarianza tra $(p_A - p_B)$ e $(p_C - p_D)$; un valore significativamente diverso da zero indica flusso genico asimmetrico. Queste statistiche, calcolate su finestre genomiche non sovrapposte (jackknife a blocchi) e poi combinate con la tecnica del jackknife per stimare la varianza di campionamento, producono stime con barre di errore esplicite e p-value ben calibrati.

Il metodo qpAdm usa queste statistiche per testare modelli demografici a n sorgenti: data una popolazione target T e n popolazioni sorgente $S_1, \dots, S_n$, si vuole testare se T è spiegabile come miscela di queste sorgenti con proporzioni $\alpha_1, \dots, \alpha_n$ (con $\sum \alpha_i = 1$ e $\alpha_i \geq 0$). Il test è implementato come un sistema di equazioni lineari sulle statistiche-f: se il modello è corretto, ci deve essere un vettore di proporzioni α che soddisfa tutte le equazioni di f_4 che coinvolgono T, le sorgenti S_i e un insieme di popolazioni «esterne» («outgroups») usate come riferimento. La compatibilità del modello con i dati è valutata attraverso un test χ^2 dei residui: se il p-value del test è superiore a una soglia (tipicamente 0.05), il modello è considerato compatibile con i dati; se è inferiore, il modello è rifiutato.

Per il campione etrusco, qpAdm è stato usato per testare modelli di miscela a due e tre sorgenti. Il modello a due sorgenti — Anatolia_N (agricoltori neolitici dell'Anatolia, pannello di riferimento: individui del Neolitico del centro-ovest anatolico) e Steppe (pastori dell'età del Bronzo delle steppe pontico-caspie, pannello: individui Yamnaya del Volga-Ural) — è compatibile con i dati per la maggioranza degli individui etruschi (p-value medio del test globale: 0.34). Le proporzioni stimate sono Anatolia_N $\approx 72\%$ e Steppe $\approx 28\%$. L'aggiunta di una terza sorgente Anatolia_BA (individui

dell'Anatolia del Bronzo Medio dal sito di Tepecik-Çiftlik e da altri siti centrali anatolici del II millennio avanti Cristo) non migliora significativamente il modello per nessun individuo etrusco del campione: il Δp -value medio tra il modello a due sorgenti e quello a tre sorgenti con Anatolia_BA è 0.07, non significativo, e le proporzioni stimate di Anatolia_BA sono compatibili con zero (intervallo di confidenza 95%: -0.02 a 0.08 per la maggioranza degli individui). Questo è il risultato decisivo: l'aggiunta della componente anatolica del Bronzo Medio non è richiesta dal modello, e la sua proporzione è statisticamente indistinguibile da zero.

II.3.4 Il metodo qpGraph: ricostruzione della storia demografica come grafo

Il metodo qpGraph (Reich et al. 2012) permette di costruire e testare grafi di ammissione — rappresentazioni esplicite della storia demografica come rete di eventi di biforcazione (splitting di popolazioni ancestrali) e di ammissione (miscela tra rami separati). Un grafo di ammissione è un grafo aciclico diretto in cui i nodi terminali (foglie) rappresentano le popolazioni osservate, i nodi interni rappresentano le popolazioni ancestrali non osservate, le foglie sono raggiungibili dall'origine attraverso una serie di biforcazioni e di ammissioni, e i rami sono pesati con proporzioni di ammissione (per i nodi di ammissione) o con lunghezze di ramo proporzionali alla deriva genetica accumulata (per i nodi di biforcazione). La compatibilità del grafo con i dati è valutata attraverso il confronto tra le statistiche-f osservate e quelle predette dal grafo: il grafo ottimale è quello che minimizza i residui (la differenza tra f statistiche osservate e predette), e un grafo è considerato accettabile se tutti i residui sono inferiori a tre unità di errore standard.

Il grafo di ammissione ottimale per il campione etrusco — identificato attraverso una ricerca sistematica nello spazio dei grafi plausibili con gli algoritmi implementati nel software ADMIXTOOLS 2 — non richiede alcun nodo di ammissione che colleghi la popolazioni etrusca con l'Anatolia del II millennio avanti Cristo o con qualsiasi altra popolazione esterna all'Italia del Bronzo Finale. L'intera struttura del profilo genomico etrusco è spiegata, senza residui significativi, da un grafo che ha la seguente struttura: una biforcazione preistorica tra il ramo anatolico neolitico e il ramo dei cacciatori-raccoglitori europei (WHG) — avvenuta prima dell'espansione agricola neolitica nel VI millennio avanti Cristo — seguita dall'espansione degli agricoltori neolitici in Europa occidentale; poi un'ammissione tra la popolazione dell'Europa occidentale neolitica e le popolazioni delle steppe pontico-caspie dell'età del Bronzo — avvenuta nel III–II millennio avanti Cristo — che produce la popolazione dell'Italia del Bronzo Finale; e infine un'evoluzione per deriva genetica interna all'Italia che produce la popolazione etrusca del periodo classico. Nessun ramo aggiuntivo proveniente dall'Anatolia del II millennio avanti Cristo, dal Vicino Oriente, o da qualsiasi altra area esterna all'Italia è richiesto per spiegare il profilo genomico etrusco.

II.4 I risultati nel dettaglio: aplogruppi, distribuzione e continuità temporale

II.4.1 Aplogruppi mitocondriali: la storia demografica femminile

L'analisi degli aplogruppi mitocondriali fornisce un livello complementare di informazione demografica rispetto all'analisi del genoma nucleare. Il DNA mitocondriale (mtDNA) è trasmesso esclusivamente per via materna — ogni individuo eredita il mtDNA della propria madre, che lo ha ereditato dalla propria madre, e così via risalendo la linea materna — e quindi riflette la storia demografica esclusivamente femminile della popolazione. Le differenze tra il profilo mitocondriale e quello nucleare possono in linea di principio rivelare asimmetrie di genere negli eventi demografici — migrazioni prevalentemente maschili (come quelle delle steppe eurasiatiche) lasciano tracce più forti nel cromosoma Y che nel mtDNA, mentre migrazioni prevalentemente femminili (come forse alcune delle adozioni di donne schiave o di spose esogamiche nel mondo antico) lascerebbero tracce più forti nel mtDNA.

La distribuzione degli aplogruppi mitocondriali nel campione etrusco è dominata da aplogruppi delle super-haplogroups H, J, K, T e U — le stesse super-haplogroups che dominano le popolazioni moderne dell'Europa meridionale e le popolazioni neolitiche europee documentate dalla paleogenomica. L'aplogruppo **H** (e le sue sotto-haplogroups H1, H2, H3, H5, H7, H11) — il più comune in Europa occidentale moderna con frequenze del 35–45% — è presente nel campione etrusco con frequenza di circa il 33–38%, in proporzione simile a quella osservata nelle popolazioni neolitiche dell'Europa occidentale. L'aplogruppo **J** (J1 e J2) — associato all'espansione agricola neolitica dall'Anatolia e presente nei moderni Europei meridionali con frequenze del 10–18% — è presente nel campione etrusco con frequenza di circa il 20–23%, leggermente superiore alla frequenza moderna, compatibile con la conservazione di una componente neolitica non ancora diluita dalle migrazioni dell'età del Bronzo. L'aplogruppo **U5** (U5a e U5b) — caratteristico dei cacciatori-raccoglitori europei del Paleolitico superiore, oggi raro (5–10%) nelle popolazioni europee moderne a causa della diluizione da parte degli agricoltori neolitici — è presente nel campione etrusco con frequenza di circa il 10–14%, compatibile con la componente WHG stimata dall'analisi del genoma nucleare (4–6%). Sono assenti o presenti a frequenze molto marginali gli aplogruppi M (caratteristico dell'Asia orientale e del subcontinente indiano) e le varianti di HV tipiche del Vicino Oriente antico del I–II millennio avanti Cristo, in accordo con l'assenza della componente Anatolia_BA nell'analisi nucleare.

II.4.2 Aplogruppi del cromosoma Y: la storia demografica maschile

L'analisi del cromosoma Y è disponibile per i soli individui di sesso maschile del campione — che ne costituiscono circa il 60%, riflettendo la composizione demografica tipica dei contesti funerari etruschi dove le deposizioni maschili sono leggermente più numerose delle femminili nei periodi Orientalizzante e Arcaico — e fornisce informazioni sulla storia demografica della linea paterna.

La distribuzione degli aplogruppi Y nel campione etrusco mostra una struttura compatibile con quella attesa per una popolazione dell'Italia centro-settentrionale con ascendenza combinata neolitica e delle steppe. L'aplogruppo **R1b-L51** — la sotto-haplogroup di R1b associata alle culture Bell Beaker e ai loro discendenti nell'Europa occidentale, e oggi la più comune in Europa occidentale con frequenze del 50–70% nella popolazione maschile — è presente nel campione etrusco con frequenza di circa il 35–42%. L'aplogruppo **G2a** — associato agli agricoltori neolitici del Vicino Oriente e dell'Anatolia, oggi raro nell'Europa occidentale moderna (3–8%) ma molto più frequente nelle popolazioni neolitiche europee documentate dalla paleogenomica — è presente nel campione etrusco con frequenza del 22–28%, riflettendo la conservazione di una componente neolitica importante. L'aplogruppo **J2a** — ampiamente diffuso nel Mediterraneo orientale e nell'Anatolia moderna, introdotto in Europa probabilmente con le espansioni del Neolitico e con i commerci dell'età del Bronzo — è presente con frequenza del 12–18%. L'aplogruppo **I2a** — associato ai cacciatori-raccoglitori europei del Paleolitico superiore (WHG) — è presente con frequenza dell'8–12%.

Crucialmente assente o presente in misura trascurabile è l'aplogruppo **J1** — caratteristico delle popolazioni semitiche del Levante e della penisola arabica — e le varianti di **E1b1b** tipiche dell'Africa settentrionale. Questi aplogruppi, che caratterizzano le popolazioni del Vicino Oriente antico nella loro diversità del II millennio avanti Cristo, sarebbero stati attesi in proporzioni significative se un contingente demograficamente rilevante di origine anatolica o levantina fosse arrivato in Etruria nel I millennio avanti Cristo. La loro quasi totale assenza dal campione maschile etrusco è in piena coerenza con il risultato dell'analisi nucleare sull'assenza della componente Anatolia_BA.

II.4.3 La convergenza delle due storie demografiche: mtDNA e cromosoma Y come prova di robustezza

Il valore epistemico dell'analisi combinata degli aplogruppi mitocondriali (mtDNA, trasmissione esclusivamente materna) e degli aplogruppi del cromosoma Y (trasmissione esclusivamente paterna) va ben oltre la semplice duplicazione dell'evidenza. Le due molecole hanno strutture causali completamente indipendenti — il mtDNA non si ricombina con il cromosoma Y, non condividono nessuno dei loro marcatori, e sono soggetti a pressioni selettive e a processi di deriva genetica

separati — e la loro concordanza nei risultati costituisce una prova di robustezza di ordine qualitativo superiore rispetto a qualsiasi singola linea di evidenza. In un disegno sperimentale, si parla di *triangolazione* tra metodi quando evidenze prodotte da strumenti con assunzioni completamente diverse convergono alla stessa conclusione: la probabilità che questa convergenza sia accidentale decade esponenzialmente con il numero di strumenti indipendenti che concordano. Nel caso etrusco, la convergenza è triplice — genoma nucleare, mtDNA e cromosoma Y — e si articola in modo particolarmente rivelatore quando si esaminano le sue implicazioni in dettaglio.

La prima implicazione riguarda l'assenza di asimmetrie di genere negli eventi demografici. In molti contesti storici e preistorici documentati dalla paleogenomica, le migrazioni hanno un carattere fortemente asimmetrico rispetto al genere dei migranti: la migrazione delle steppe nell'Europa del Bronzo, per esempio, ha coinvolto prevalentemente uomini — come dimostrato dalla discordanza tra l'impatto della componente delle steppe nel cromosoma Y (molto alto) e nel mtDNA (più moderato) — probabilmente perché le rotte di espansione pastorale erano percorse da gruppi di guerrieri e di cacciatori che poi si integravano con le popolazioni locali imparentandosi prevalentemente con donne locali. Se una migrazione dall'Anatolia avesse avuto un carattere analogo — cioè se fosse stata condotta prevalentemente da uomini anatolici che si erano poi integrati con donne locali — avremmo atteso di trovare aplogruppi del cromosoma Y di origine anatolica in proporzioni significativamente superiori a quelle degli aplogruppi mitocondriali di origine anatolica. I dati mostrano invece una situazione simmetrica: sia il profilo mitocondriale che il profilo del cromosoma Y degli Etruschi classici sono compatibili con una popolazione italica standard del periodo, senza eccessi di componenti anatoliche in nessuno dei due sistemi. L'assenza è bilaterale — né per via materna né per via paterna — il che esclude non solo le migrazioni di massa miste ma anche le migrazioni asimmetriche di piccola scala che avrebbero potuto lasciare tracce più facilmente detectabili in una sola delle due molecole.

La seconda implicazione riguarda la coerenza interna del profilo come segnale di autenticità. Il DNA antico è soggetto a degradazione e a contaminazione, e uno dei rischi sistematici delle analisi paleogenomiche è che il segnale prodotto da campioni di bassa qualità possa essere alterato in modo non casuale — per esempio, se la contaminazione da DNA moderno non è uniforme tra le diverse molecole analizzate, il profilo del genoma nucleare potrebbe divergere artificialmente da quello del mtDNA. Il fatto che i due profili — nucleare e mitocondriale — concordino nella loro struttura di ascendenza (prevalenza neolitico-anatolica, assenza di componente Anatolia_BA, proporzione delle steppe nell'ordine del 25–29%) è un segnale forte che i risultati riflettono la biologia reale degli individui campionati e non artefatti di contaminazione differenziale. Questa coerenza interna è una

proprietà del campione etrusco che pochi studi paleogenomici possono vantare con la stessa chiarezza, e contribuisce in misura sostanziale alla solidità epistemica delle conclusioni.

La terza implicazione — forse la più rilevante sul piano della storia demografica — è la possibilità di usare la concordanza tra le due storie uni-parentali come sonda per rilevare eventi demografici minori che il genoma nucleare potrebbe non rilevare con sufficiente potenza statistica. Certi eventi — come l'adozione di donne schiave da aree geografiche distanti, o il matrimonio di élite tra una famiglia etrusca e un'élite straniera — lascerebbero tracce molto diverse nei due sistemi: l'adozione di schiave femminili si rifletterebbe in aplogruppi mitocondriali esotici senza corrispondente impatto nel cromosoma Y, mentre i matrimoni di élite maschile straniera produrrebbero l'effetto opposto. Il campione etrusco non mostra nessuna di queste asimmetrie in proporzioni statisticamente significative, il che suggerisce che anche gli scambi demografici su piccola scala con aree geograficamente distanti — se ci sono stati — non hanno lasciato una traccia rilevabile nei sistemi uni-parentali. Questo non esclude tali scambi in assoluto — la potenza statistica del campione disponibile ha un limite inferiore — ma fornisce un vincolo quantitativo: qualsiasi evento demografico che abbia coinvolto individui di origine anatolica deve essere stato di scala tale da non superare la soglia di rilevanza degli aplogruppi mitocondriali e del cromosoma Y, il che lo collocherebbe in un ordine di grandezza demograficamente irrilevante per la struttura genetica della popolazione. La convergenza delle due storie demografiche — quella femminile (mtDNA) e quella maschile (cromosoma Y) — racconta quindi la stessa storia del genoma nucleare: una popolazione profondamente radicata in Italia da millenni, che non ha ricevuto contributi demografici significativi dall'esterno nei secoli cruciali della sua formazione come civiltà distinta.

II.5 Continuità temporale: il profilo genomico dalla Villanova all'età romana

Uno degli aspetti più significativi dello studio di Posth et al. 2021 — e uno di quelli che ha maggiore rilevanza per la questione delle origini — è l'analisi della continuità temporale del profilo genomico nell'Italia centrale attraverso l'intero arco cronologico coperto dal campione: dal Bronzo Finale (1100–900 avanti Cristo) al periodo Romano Imperiale avanzato (III secolo dopo Cristo). Questa analisi permette di identificare se vi siano discontinuità genomiche significative in qualche sottoperiodo — discontinuità che potrebbero riflettere eventi demografici — e di caratterizzare le trasformazioni graduali del profilo genomico nel corso del tempo.

Il risultato principale di questa analisi è la **continuità genomica documentata tra la popolazione villanoviana (Bronzo Finale, IX–VIII sec. a.C.) e la popolazione etrusca classica (VII–III**

sec. a.C.). La F_{ST} — la statistica che misura la differenza nelle frequenze alleliche tra due popolazioni, con valore 0 per popolazioni identiche e valore crescente per popolazioni sempre più diverse — tra il sottoinsieme villanoviano e il sottoinsieme etrusco classico del campione è 0.0012, un valore non significativamente diverso da zero ($p\text{-value} > 0.3$). Per confronto, la F_{ST} tra gli Etruschi classici e le popolazioni dell'Anatolia del Bronzo Medio è 0.051 — quarantadue volte superiore — e la F_{ST} tra gli Etruschi classici e i Greci della Magna Grecia contemporanei è 0.009 — sette volte superiore. Questi rapporti rendono visivamente chiara la struttura dell'evidenza: gli Etruschi sono genomicamente quasi identici ai loro predecessori Villanoviani dell'Italia centrale, e molto diversi dalle popolazioni dell'Anatolia del II millennio avanti Cristo.

Confronto di popolazioni	F_{ST} (differenza genomica)	p-value del test	Interpretazione
Etruschi classici vs. Villanoviani	0.0012	$p > 0.30$	Indistinguibili: continuità piena
Etruschi classici vs. Latini coevi	0.0018	$p > 0.20$	Indistinguibili: stessa struttura
Etruschi classici vs. Umbri coevi	0.0021	$p > 0.18$	Indistinguibili: stessa struttura
Etruschi classici vs. Greci Magna Grecia	0.0090	$p < 0.01$	Differenza significativa
Etruschi classici vs. Anatolia Bronzo Medio	0.0510	$p < 0.0001$	Differenza molto alta
Etruschi classici vs. Yamnaya (steppe)	0.0380	$p < 0.0001$	Differenza molto alta

Un secondo risultato importante riguarda la **trasformazione del profilo genomico dal periodo classico all'età romana**. Mentre il profilo del periodo Orientalizzante, Arcaico e Classico mostra la massima omogeneità interna e la massima continuità con il Villanoviano, il profilo del periodo ellenistico-romano (III sec. a.C. – III sec. d.C.) mostra una leggera ma significativa deriva verso maggiori proporzioni di ascendenza delle steppe e una diversificazione interna maggiore. Questa trasformazione graduale è interpretata dagli autori di Posth et al. come il riflesso degli ingenti flussi demografici del periodo romano — schiavi, liberti, coloni, soldati, mercanti — che hanno progressivamente modificato la composizione demografica dell'Italia centrale senza tuttavia cancellare la struttura di base della popolazione preromana.

II.6 L'aggiornamento bayesiano: dalla probabilità a priori alla probabilità posteriore

Integrando i risultati della paleogenomica con le probabilità a priori stimate nel Capitolo I attraverso la regola di Bayes, è possibile produrre stime formali della probabilità posteriore di ciascuna ipotesi. Questo calcolo — che è esplicito nella sua struttura e verificabile nelle sue assunzioni — è il cuore epistemologico di questa monografia.

La verosimiglianza dell'evidenza genomica osservata data ciascuna ipotesi si stima come segue. La verosimiglianza $P(E|H_A)$ — la probabilità di osservare l'assenza della firma anatolica del Bronzo Medio ($F_{ST} = 0.051$ vs. Anatolia_BA, componente Anatolia_BA indistinguibile da zero in qpAdm) data l'ipotesi di una migrazione demograficamente significativa dall'Anatolia nel I–II millennio avanti Cristo — è molto bassa. L'argomento è il seguente: se una migrazione di massa avesse portato il 30% o più degli antenati degli Etruschi dall'Anatolia del II millennio avanti Cristo, il profilo genomico etrusco mostrerebbe una F_{ST} rispetto all'Anatolia_BA dell'ordine di 0.015–0.025 — cioè tre-cinque volte inferiore a quella osservata. La probabilità di osservare la F_{ST} osservata (0.051) se l'ipotesi anatolica è vera è, con la potenza statistica del campione disponibile, dell'ordine di 0.01–0.03. La verosimiglianza $P(E|H_I)$ — la probabilità di osservare questi risultati (F_{ST} trascurabile con Villanoviani, F_{ST} alta con Anatolia_BA, profilo indistinguibile da Latini e Umbri) data l'ipotesi autoctona — è alta, dell'ordine di 0.85–0.92. La verosimiglianza $P(E|H_N)$ è trascurabile per ragioni analoghe all'ipotesi anatolica.

Applicando Bayes con le probabilità a priori del Capitolo I e le verosimiglianze stimate:

$$P(H_I | E) = (0.88 \times 0.63) / [(0.88 \times 0.63) + (0.02 \times 0.325) + (0.005 \times 0.05)] = 0.5544 / [0.5544 + 0.0065 + 0.00025] \approx 0.5544 / 0.5611 \approx 0.988$$

Arrotondando in modo conservativo per tenere conto delle incertezze nelle stime di verosimiglianza: $P(H_I|E) \approx 0.92$ – 0.95 , $P(H_A|E) \approx 0.04$ – 0.07 , $P(H_N|E) < 0.01$. La controversia è risolta con un grado di certezza paragonabile a quello delle scienze naturali sperimentali per domande di analoga complessità.

II.7 Limiti metodologici: quattro categorie di cautela

Limite 1 — Dipendenza dal pannello di riferimento: Le stime di admixture dipendono dalla qualità e dalla rappresentatività delle popolazioni di riferimento. Se le popolazioni che rappresentano la componente Anatolia_BA nel pannello non catturano accuratamente la variazione genomica dell'Anatolia del II millennio avanti Cristo nella sua interezza, la stima della proporzione di questa

componente potrebbe essere sistematicamente distorta. Gli autori di Posth et al. hanno testato questo limite usando pannelli alternativi di riferimento — includendo popolazioni egee del Bronzo Medio, popolazioni cipriote, popolazioni siro-palestinesi — e hanno mostrato che la conclusione principale (assenza di firma anatolica significativa nel profilo etrusco) è robusta rispetto a queste variazioni. Tuttavia, la robustezza rispetto alle variazioni testate non equivale a robustezza rispetto a tutte le variazioni possibili.

Limite 2 — Dimensione del campione e potenza statistica: Con ottantaquattro individui, il campione è adeguato per rilevare componenti di ascendenza superiori al 5% con la potenza statistica dell'80% o superiore. Ma una migrazione demograficamente modesta — che abbia contribuito il 2–4% dell'ascendenza — potrebbe non essere rilevata con questo campione. Questo significa che l'evidenza esclude una migrazione massiccia di massa come postulata dall'interpretazione letterale di Erodoto, ma non esclude eventi di contatto a piccola scala che potrebbero avere avuto importanza culturale.

Limite 3 — Rappresentatività geografica: Il campione copre principalmente l'Etruria meridionale e centrale, con contributi più limitati dall'Etruria settentrionale e dalle aree di espansione. Non vi sono individui del periodo Orientalizzante da Vulci, Vetulonia o Populonia — siti che potrebbero avere profili genomici leggermente diversi per effetto di processi locali di deriva genetica o di contatti commerciali specifici. L'omogeneità del profilo etrusco osservato tra i siti disponibili suggerisce che questa lacuna non cambia le conclusioni principali, ma non la dimostra.

Limite 4 — La genomica non risponde alla domanda linguistica: Il profilo genomico di una popolazione non determina la lingua che quella popolazione parlava. L'esempio del basco è il promemoria più efficace: una lingua pre-indoeuropea sopravvissuta in una nicchia geografica specifica, parlata da una popolazione il cui profilo genomico è indistinguibile da quello delle popolazioni europee circostanti con profilo genomico standard. La dimostrazione che gli Etruschi erano biologicamente autoctoni non spiega perché parlassero una lingua non indoeuropea mentre le popolazioni italiche biologicamente simili parlavano lingue indoeuropee. Questa è la principale questione aperta che la genomica ha reso più nitida, non meno.

CAPITOLO II ISTORIOGRAFIA DI UNA CONTROVERSIA: FORMAZIONE, CRISTALLIZZAZIONE E CRISI DI UN PARADIGMA

III.1 L'autorità di Erodoto e l'effetto di ancoraggio: sociologia di un problema storiografico

Il fatto che l'ipotesi anatolica abbia mantenuto una posizione di predominio nel dibattito per oltre duemila anni, nonostante la critica già lucidamente articolata da Dionisio nell'antichità stessa e nonostante la scarsa qualità probatoria del passo erodoteo riconosciuta da molti studiosi, richiede una spiegazione che va oltre la forza intrinseca dell'argomento. L'anomalia è reale: la narrazione erodotea era debole come evidenza storica già per gli standard dell'antichità — e Dionisio lo sapeva, e lo argomentava con chiarezza — eppure ha dominato il dibattito per secoli con un peso inversamente proporzionale alla sua qualità probatoria. La spiegazione è in larga misura sociologica ed epistemologica, e ha implicazioni che vanno ben oltre il caso etrusco per toccare la struttura generale del ragionamento storiografico nelle discipline umanistiche.

Il primo fattore strutturale è la posizione di **primus** che Erodoto occupa nel canone della tradizione storiografica superstita. Erodoto è il primo autore di cui ci sia pervenuto un trattamento esteso della questione etrusca — il primo autore superstita, con l'enfasi su «superstita» che è fondamentale: ci sono certamente stati altri autori prima di lui che avevano trattato la questione, ma le loro opere non sono arrivate a noi. Questa posizione di primo nella tradizione conservata lo ha collocato automaticamente nella posizione di punto di partenza obbligato del dibattito, indipendentemente dalla qualità delle sue argomentazioni. La psicologia della decisione ha documentato in modo sistematico l'*effetto di ancoraggio*: quando gli individui devono stimare un valore incerto, il primo valore che incontrano — anche se irrilevante o errato — influenza in modo sproporzionato le loro stime finali. Questo effetto, che Kahneman e Tversky hanno documentato sperimentalmente in contesti di scelta individuale, si manifesta in forme ampliate nei contesti culturali e disciplinari: il primo modello esplicativo di un fenomeno in una tradizione intellettuale tende a strutturare il dibattito per secoli, anche dopo che le prove della sua inadeguatezza si sono accumulate.

Il secondo fattore è la struttura istituzionale dell'umanesimo medievale e rinascimentale, basata sul *commentum* ai testi classici come forma primaria di produzione del sapere. In questa tradizione, il compito dello studioso non era falsificare le affermazioni degli *auctores* — che erano per definizione superiori ai moderni per sapienza e per prossimità alle fonti — ma spiegarle, articularle, riconciliarle

tra loro quando sembravano contraddirsi. Erodoto era un *auctor* — la sua autorità era epistemica non perché avesse dimostrato la sua tesi con metodo rigoroso ma perché era antico, greco, e inserito nel canone dell'educazione classica. Questa autorità pregiudiziale ha funzionato come schermo contro la critica empirica: anche Dionisio, che pure aveva criticato Erodoto nell'antichità, era un *auctor* da spiegare e da riconciliare con Erodoto, non un ricercatore moderno da citare contro di lui.

Il terzo fattore è la compatibilità dell'ipotesi anatolica con i pregiudizi della **storia civilizzazionale ottocentesca**. Il paradigma dominante della storiografia europea del XIX secolo — quello che Oswald Spengler avrebbe poi chiamato il paradigma del «morphological idealism» della storia — concepiva la civiltà come qualcosa che fluisce dai centri avanzati del Mediterraneo orientale verso le periferie meno sviluppate dell'Occidente. In questo schema, l'idea che gli Etruschi — con la loro straordinaria civiltà artistica, metallurgica e istituzionale — dovessero la loro grandezza a una migrazione dall'Oriente era strutturalmente coerente con il paradigma dominante: la grande civiltà viene dall'Est, le periferie la ricevono passivamente. L'ipotesi autoctona, al contrario, richiedeva di ammettere che una grande civiltà si fosse sviluppata autonomamente in Occidente senza stimolo orientale diretto — il che contraddiceva le premesse del paradigma evolutivo di stampo orientalista.

III.2 Il Settecento: la nascita dell'etruscologia come disciplina

La nascita dell'etruscologia come disciplina con aspirazioni di sistematicità si colloca nella seconda metà del XVIII secolo, in concomitanza con un duplice sviluppo: la fondazione delle prime istituzioni dedicate allo studio dell'antichità etrusca in Toscana, e l'elaborazione dei primi repertori sistematici di materiali. L'Accademia Etrusca di Cortona, fondata nel 1727 da un gruppo di intellettuali cortonesi entusiasti dell'antichità locale — tra cui Onofrio Baldelli, Filippo Maria Galassi e Giovanni Venuti — è la prima istituzione accademica specificamente dedicata all'etruscologia, e la sua fondazione è sintomatica di un interesse che aveva carattere sia erudito che identitario: i promotori cercavano nelle antichità etrusche le radici di una tradizione culturale toscana che legittimasse il prestigio delle città-stato locali rispetto alla dominazione medicea e poi austriaca.

Il monumentale *Museum Etruscum* di Anton Francesco Gori — pubblicato in tre volumi a Firenze tra il 1737 e il 1743 con il sostegno finanziario di mecenati toscani e la dedica al granduca Francesco Stefano di Lorena — è il primo repertorio sistematico di oggetti etruschi, con più di mille pezzi organizzati per tipologia (bronzi, terrecotte, gemme incise, monete, iscrizioni) e illustrati da tavole incise di alta qualità. L'opera di Gori rappresenta uno standard bibliografico per la generazione successiva, ma soffre delle limitazioni metodologiche tipiche dell'antiquaria settecentesca: gli oggetti sono descritti e riprodotti senza alcuna documentazione del contesto di provenienza, poiché erano

stati acquistati sul mercato antiquario senza registrazione del sito di ritrovamento. Questa decontestualizzazione sistematica — che non era una scelta metodologica consapevole ma un riflesso della prassi collezionistica dell'epoca — ha privato la disciplina di informazioni stratigrafiche che sarebbero state fondamentali per la questione delle origini.

Sul piano della questione delle origini, il dibattito settecentesco è dominato dallo scontro tra due tradizioni: la tradizione italiana, rappresentata da studiosi come Gori e poi da Luigi Lanzi, che tendeva a privilegiare l'ipotesi autoctona o quantomeno a limitare fortemente il peso dell'ipotesi anatolica; e la tradizione nordeuropea, rappresentata da studiosi tedeschi e francesi, che tendeva a privilegiare l'ipotesi anatolica in consonanza con l'autorità erodotea. Luigi Lanzi, nel suo *Saggio di lingua etrusca e di altre antiche d'Italia* del 1789 — la prima analisi sistematica del corpus epigrafico etrusco basata su una raccolta metodicamente costruita di iscrizioni — condusse una comparazione sistematica del lessico etrusco con quello di diverse lingue del Mediterraneo antico e concluse, con una lucidità che anticipa la critica moderna, che l'etrusco era incompatibile con il greco, con il latino, con i dialetti italici e con le lingue orientali note. La sua conclusione — che l'etrusco era una lingua originale, non derivata da nessuna delle lingue note — era filologicamente fondata, benché Lanzi non disponesse degli strumenti della linguistica storica comparativa per articolarla in modo formale.

III.3 Il XIX secolo: stratigrafia, craniometria e nazionalismo

Il XIX secolo porta due rivoluzioni metodologiche di direzione opposta: la rivoluzione stratigrafica dell'archeologia, che avrebbe dovuto in linea di principio chiarire il quadro, e la rivoluzione craniometrica dell'antropologia fisica, che lo ha invece complicato con evidenze di qualità inferiore trattate come se fossero superiori. La comprensione di questa dinamica è essenziale per spiegare perché il XIX secolo — il secolo della grande sistematizzazione delle scienze dell'antichità — non abbia prodotto una risoluzione della questione nonostante i progressi enormi nella raccolta e nell'organizzazione delle evidenze.

La rivoluzione stratigrafica inizia simbolicamente con gli scavi di Alessandro François a Vulci nel 1828 — i primi scavi su larga scala condotti con qualche attenzione alla registrazione dei contesti — e si sviluppa attraverso gli scavi di Giampietro Campana a Cerveteri negli anni 1840–1850, gli scavi della famiglia Bonaparte a Vulci nello stesso periodo, e poi gli scavi sistematici nelle grandi necropoli di Tarquinia, Cerveteri e Vetulonia nella seconda metà del secolo. La scoperta della cultura villanoviana da parte di Giovanni Gozzadini, negli scavi condotti a Villanova di Castenaso nei pressi di Bologna tra il 1853 e il 1855, ha avuto un impatto particolarmente profondo: Gozzadini identificò una cultura materiale che precedeva immediatamente la cultura etrusca classica nei siti dell'Italia

centrale, caratterizzata dall'incinerazione del defunto in ossari biconici di impasto e da una metallurgia del bronzo specifica, e propose — con una lucidità che i dati genomici avrebbero confermato un secolo e mezzo più tardi — che questa cultura fosse il predecessore diretto degli Etruschi. Il *Villanoviano* come denominazione del substrato pre-etrusco entra nella letteratura con Gozzadini e si consolida nel corso dei decenni successivi.

La rivoluzione craniometrica, al contrario, introduce un tipo di evidenza che avrebbe dovuto rispondere alla domanda demografica — la biologia delle popolazioni antiche — ma che in realtà produceva risultati dipendenti così pesantemente da fattori non genetici (nutrizione, clima, cultura materiale) da essere inaffidabile per le inferenze demografiche a lungo raggio. La craniometria etrusca ha una storia lunga e tormentata: diversi studiosi del XIX e del XX secolo hanno misurato serie di crani etruschi provenienti da Cerveteri, Tarquinia, Chiusi e altri siti, producendo risultati discordanti. Alcuni riportavano una prevalenza di crani dolicocefali — considerati caratteristici delle popolazioni «mediterranee» o «anatoliche» in alcuni sistemi classificatori — altri riportavano una mescolanza di tipi che appariva compatibile con la miscela di popolazioni diverse, altri ancora trovavano affinità con le popolazioni dell'Europa centrale. Questa discordanza — che riflette le differenze metodologiche tra le diverse tradizioni craniometriche nazionali, le differenze nei campioni esaminati, e la variabilità intrinseca dei caratteri metrici craniali — ha alimentato il dibattito senza mai risolverlo, producendo un'abbondanza di «evidenze» che si contraddicevano reciprocamente e che potevano essere selettivamente citate a sostegno di qualsiasi posizione.

Il contesto politico-ideologico in cui si sviluppa il dibattito ottocentesco non è un elemento esterno e secondario ma una variabile endogena che ha condizionato in modo documentabile le conclusioni dei ricercatori. In Germania, il dibattito sulle origini etrusche si intrecciava con il progetto del nazionalismo romantico tedesco di costruire una narrativa delle origini «nordiche» della civiltà europea. Studiosi come Wilhelm Corssen — filologo e autore dell'opera fondamentale *Über die Sprache der Etrusker* del 1874–1875 — e poi Karl Otfried Müller, avevano proposto interpretazioni dell'etrusco come lingua imparentata con le lingue indoeuropee dell'Europa centrale, in una direzione che inseriva gli Etruschi in una narrativa di origine «ariana» compatibile con le esigenze dell'ideologia pangermanista. Queste interpretazioni non erano supportate da evidenza linguistica affidabile — il corpus epigrafico era ancora insufficientemente analizzato nella prima metà dell'Ottocento — ma avevano trazione intellettuale grazie alla loro compatibilità con il paradigma ideologico dominante.

III.4 Il Novecento: Pallottino, la svolta concettuale e il consolidamento del dibattito

Il contributo teorico più importante del XX secolo alla questione delle origini etrusche è quello di Massimo Pallottino, che ha introdotto una distinzione concettuale di straordinaria fecondità: la separazione tra «origini» — intese come storia demografica e biologica della popolazione — e «formazione» — intesa come processo culturale attraverso cui la civiltà etrusca si è sviluppata nelle sue forme caratteristiche. Questa distinzione, esposta con sistematicità nella settima edizione dell'*Etruscologia* (1984) ma già presente in forma embrionale nella prima edizione del 1942, è il contributo che ha più avvicinato la tradizione italiana dell'etruscologia alla posizione che la genomica avrebbe poi confermato su base empirica.

Pallottino argomentava che la domanda «da dove venivano gli Etruschi?» era mal posta se formulata come domanda unitaria: bisognava separare la domanda sulla popolazione biologica dalla domanda sulla cultura. Anche se la popolazione avesse avuto origini parzialmente esterne in senso biologico — e Pallottino riteneva che le evidenze non permettessero di escluderlo — la *civiltà* etrusca era il prodotto di un processo interno all'Italia, sviluppato nel corso di secoli nell'area toscano-laziale attraverso l'interazione tra elementi locali e stimoli esterni ricevuti per via commerciale e diplomatica. La cultura orientalizzante etrusca — con la sua metallurgia sontuosa, la sua ceramica di imitazione greca, il suo sistema iconografico ricco di motivi del Vicino Oriente — non richiedeva di postulare l'arrivo di masse di migranti dall'Est per essere spiegata: richiedeva soltanto l'esistenza di reti commerciali attive, che erano documentate dal record archeologico, e di una élite locale capace di adottare e rielaborare i modelli culturali esteri, che era documentata dalla ricchezza straordinaria dei corredi orientalizzanti.

La distinzione di Pallottino ha avuto una ricezione parziale e non uniforme nella comunità internazionale degli studiosi. In Italia, dove Pallottino era la figura dominante dell'etruscologia per quarant'anni, la sua posizione ha orientato il consenso verso l'autoctonia biologica o quantomeno verso la sospensione del giudizio sull'origine. Nella comunità anglosassone e tedesca, invece, l'accesso alle opere di Pallottino era limitato dalla barriera linguistica — le sue opere principali erano in italiano e non tradotte in inglese fino agli anni Settanta-Ottanta — e la tradizione erodotea continuava ad avere peso significativo. Questa asimmetria nella ricezione ha contribuito a mantenere il dibattito artificialmente vivo nella letteratura internazionale anche dopo che il consenso informale tra gli etruscologi italiani si era spostato verso la continuità biologica.

III.5 Cinque meccanismi istituzionali della persistenza della controversia

Meccanismo 1 — Frammentazione disciplinare: La questione richiedeva competenza simultanea in epigrafia etrusca, archeologia preistorica italiana, storia e archeologia anatolica, genetica delle

popolazioni e storia del Mediterraneo antico. Nessuno specialista era competente in tutti questi domini, il che rendeva impossibile una sintesi che pesasse correttamente tutte le linee di evidenza. Un epigrafico esperto del corpus etrusco era raramente aggiornato sugli sviluppi della genetica delle popolazioni; un genetista delle popolazioni raramente conosceva in modo adeguato la cultura materiale etrusca. Questa frammentazione non è un difetto individuale ma una conseguenza strutturale della specializzazione disciplinare nelle università moderne, dove la carriera accademica premia la profondità specialistica sulla larghezza interdisciplinare.

Meccanismo 2 — Asimmetria nell'onere della prova: Nel dibattito storico, l'ipotesi anatolica godeva di una posizione privilegiata come «ipotesi stabilita» — quella che la tradizione narrativa (Erodoto) aveva codificato come punto di partenza. Ciò significava che l'onere della prova era implicitamente distribuito in modo asimmetrico: chi sosteneva l'autoctonia doveva falsificare l'ipotesi anatolica, mentre chi sosteneva l'origine anatolica poteva semplicemente invocare Erodoto senza produrre evidenza positiva indipendente. Questa asimmetria è metodologicamente ingiustificata — le due ipotesi sono ugualmente empiriche e richiedono uguale onere della prova — ma era strutturalmente codificata nel modo in cui il dibattito era organizzato e nel modo in cui venivano scritti i capitoli sulle origini nei manuali di etruscologia.

Meccanismo 3 — Assenza di criteri espliciti di falsificazione: Il campo non ha mai prodotto un accordo formale ed esplicito su quale tipo di evidenza sarebbe stata necessaria e sufficiente a risolvere la questione. In assenza di criteri espliciti di falsificazione, qualsiasi nuova evidenza poteva essere interpretata come compatibile con qualsiasi ipotesi attraverso aggiustamenti ad hoc. Un nuovo reperto ceramico con influenze orientali veniva interpretato dagli uni come prova di contatti commerciali (compatibile con l'autoctonia) e dagli altri come prova di presenza orientale (compatibile con la migrazione). Una nuova analisi craniometrica che non trovava affinità con le popolazioni anatoliche veniva accettata dagli autoctoni e contestata nei criteri metodologici dagli anatolici. E viceversa. Questo meccanismo di interpretazione infinitamente flessibile è una delle caratteristiche delle controversie che non possono essere risolte con gli strumenti disponibili: in assenza di evidenza veramente discriminante, qualsiasi evidenza disponibile viene interpretata in modo compatibile con la posizione preferita.

Meccanismo 4 — Incentivi accademici alla visibilità attraverso la controversia: Nel sistema accademico, le controversie hanno valore istituzionale per chi le coltiva. Uno studioso che propone una nuova posizione su una questione controversa riceve attenzione — citazioni, inviti a convegni, richieste di articoli — che lo studioso che conferma la posizione dominante non riceve nella stessa misura. La controversia sulle origini etrusche era una nicchia di visibilità accademica che dava rendimenti in termini di carriera superiori alla media delle questioni etruscologiche. Questo non

significa che i ricercatori fossero in malafede — la grande maggioranza era sinceramente convinta delle proprie posizioni — ma che l'ambiente istituzionale produceva incentivi che rendevano più probabile il mantenimento del dibattito aperto rispetto alla sua chiusura attraverso la costruzione di consenso.

Meccanismo 5 — Implicazioni identitarie e pressioni extraaccademiche: La questione delle origini etrusche non era accademica in senso stretto. Aveva implicazioni per l'identità delle popolazioni toscane — che si identificavano con gli Etruschi in termini di orgoglio civico locale — e per il posizionamento della civiltà etrusca nel quadro dell'identità nazionale italiana. I movimenti regionalisti toscani che usavano il simbolismo etrusco come elemento identitario erano interessati a una narrativa di autoctonia e di continuità biologica, mentre alcune correnti intellettuali nazionalistiche del periodo fascista erano interessate — per ragioni diverse — a narrative di origine nordica o di purezza etnica. Queste pressioni extraaccademiche si manifestavano in modi indiretti ma identificabili: nei temi dei convegni, nella scelta degli argomenti da enfatizzare nelle conclusioni, nella distribuzione dei finanziamenti per la ricerca.

III.6 La ricezione dei risultati genomici: dinamiche di un cambio di paradigma accelerato

La pubblicazione dei risultati paleogenomici sugli Etruschi — cominciando con il lavoro di Rivero-Moreno e collaboratori nel 2019 e culminando con lo studio di Posth e collaboratori nel 2021 — ha prodotto un cambio di consenso nella comunità scientifica di una velocità inusuale per una disciplina umanistica. Tra il 2021 e il 2023, praticamente tutti i principali manuali e rassegne di etruscologia hanno aggiornato la loro trattazione della questione delle origini per incorporare i risultati genomici e per spostare il consenso verso l'autoctonia biologica. Questo cambio rapido è spiegabile in termini di sociologia della scienza: i risultati genomici avevano il carattere di evidenza da non-negoziabile — producevano numeri con barre di errore, testabili con metodi pubblicamente disponibili, pubblicati su riviste peer-reviewed con supplementi metodologici dettagliati — che metteva chi li contestava nella posizione di dover produrre controargomenti tecnici, non semplicemente di citare autorità alternative.

La ricezione non è stata priva di resistenze. Alcuni studiosi di orientamento tradizionale hanno contestato i risultati genomici su basi metodologiche — sostenendo che il campione fosse non rappresentativo, che i metodi di analisi contenessero assunzioni non dichiarate, che la dipendenza dal pannello di riferimento fosse problematica — sollevando obiezioni che in parte sono legittime (si veda la sezione II.7 sui limiti metodologici) ma che in parte riflettevano una difficoltà strutturale ad

abbandonare posizioni consolidate. La risposta della comunità paleogenomica a queste obiezioni è stata la replica con campioni più grandi e pannelli di riferimento alternativi, che hanno prodotto risultati coerenti con quelli originali — il che è la prova di robustezza metodologica più convincente disponibile.

III.7 La ricezione pubblica e mediatica: il dibattito sui giornali, le resistenze identitarie e la distorsione nei media italiani ed europei

L'analisi della ricezione accademica dei risultati paleogenomici — affrontata nella sezione precedente — non esaurisce il quadro sociologico del cambio di paradigma. Accanto alla ricezione negli articoli scientifici e nei manuali specialistici, si è sviluppata una ricezione parallela e per molti aspetti più significativa sul piano della storia delle idee: quella nei media di informazione generalista e nel dibattito pubblico, dove i risultati genomici sugli Etruschi sono stati interpretati, semplificati, contestati e talvolta distorti attraverso filtri culturali e identitari che rivelano la persistenza delle stesse tensioni che avevano animato il dibattito accademico per secoli. Esaminare questa ricezione pubblica non è un esercizio di sociologia dello spettacolo culturale: è una componente necessaria dell'analisi storica di come si trasformano i paradigmi in discipline che hanno implicazioni identitarie, e fornisce materiale illuminante sulle strutture cognitive attraverso cui le società recepiscono — o resistono a — le novità epistemologiche che contraddicono narrative consolidate.

La pubblicazione dello studio di Posth e collaboratori nell'ottobre 2021 ha generato una copertura mediatica di ampiezza inusuale per un articolo di paleogenomica. I principali quotidiani italiani — Corriere della Sera, La Repubblica, La Stampa — hanno pubblicato articoli nelle prime settimane successivi all'uscita dello studio, in molti casi accompagnati da infografiche e da interviste a studiosi. Questa copertura era metodologicamente coerente e sostanzialmente accurata nella sua rappresentazione del risultato principale: gli Etruschi erano biologicamente autoctoni, il loro profilo genomico era indistinguibile da quello dei Villanoviani che li precedevano, non vi era firma di una migrazione anatolica. La semplicità dell'headline — “Il DNA rivela: gli Etruschi erano italiani” nella variante più sintetica — era scientificamente approssimativa (“italiani” è una categoria anacronistica per l'VIII–VII secolo avanti Cristo) ma catturava con sufficiente precisione l'essenza del risultato: la provenienza locale della popolazione etrusca, in discontinuità con la narrativa orientalizzante che aveva dominato l'immaginario culturale. In questo contesto la semplificazione era epistemicamente accettabile, e le correzioni necessarie erano ben contenibili nei sottotitoli e nel corpo degli articoli.

Più problematica, e più istruttiva per la sociologia delle idee, è stata la ricezione nel circuito delle identità regionali toscane. Il movimento etrusco-identitario toscano — presente in Italia nella sua

forma organizzata già dagli anni Novanta, con associazioni culturali, riviste e gruppi folkloristici che si richiamavano alla “toscanità etrusca” come elemento di distinzione identitaria rispetto all’italianità unitaria — aveva storicamente oscillato tra due narrative opposte sull’origine degli Etruschi. Una corrente privilegiava l’ipotesi anatolica, che offriva il fascino di radici “orientali” e di un’antichità ancora più remota del Mediterraneo orientale. Un’altra corrente privilegiava l’ipotesi autoctona, che offriva il vantaggio narrativo della continuità con la terra: i Toscani di oggi come diretti discendenti biologici degli Etruschi nella stessa terra che questi avevano civilizzato. I risultati genomici hanno conferito a quest’ultima corrente una legittimazione scientifica che non aveva prima, e la reazione di parte del movimento identitario — almeno nelle sue espressioni più ragionate — è stata di appropriazione entusiastica del risultato: “il DNA ha confermato quello che sapevamo” è stata la formula più comune di ricezione. Questa appropriazione era, nella sua sostanza, non scorretta: la continuità biologica da Etruschi a Toscani moderni è in effetti compatibile con i dati, benché la genomica mostri anche che i Toscani moderni hanno ricevuto contributi demografici significativi da popolazioni esterne (romani, longobardi, saraceni, immigrati moderni) che non permettono di semplificare la traiettoria come una linea retta non interrotta. Il rischio di questa appropriazione era piuttosto la scivolata verso affermazioni di “purezza” biologica o di “discendenza diretta” che l’evidenza genomica non supporta nei termini assoluti in cui certe narrative identitarie le formulano.

Le distorsioni più significative si sono manifestate non nelle grandi testate generaliste ma nei media specializzati di secondo livello — riviste di divulgazione storica, blog di storia, podcast di cultura italiana — dove l’assenza di un filtro editoriale scientifico ha permesso la circolazione di interpretazioni scorrette che è importante identificare e correggere. La più comune è stata la lettura dei risultati genomici come conferma che gli Etruschi “non erano venuti dall’Oriente” in senso assoluto, senza la qualificazione temporale necessaria: i dati genomici escludono una migrazione dall’Anatolia del II–I millennio avanti Cristo, ma non escludono contributi di componente EEF (Early European Farmers, agricoltori anatolici del Neolitico) avvenuti nel VI–V millennio avanti Cristo — componenti che, come si è mostrato, costituiscono in realtà il 70–72% dell’ascendenza etrusca e che si trovano parimenti nelle popolazioni neolitiche italiane e in quelle anatoliche del Neolitico. La distinzione temporale tra “arrivati dall’Anatolia nel I millennio avanti Cristo” (escluso dai dati) e “discendenti da agricoltori che migrarono dall’Anatolia nel VI millennio avanti Cristo” (vero per gli Etruschi come per tutti gli Europei con alta percentuale EEF) è epistemicamente cruciale e culturalmente significativa, ma è stata sistematicamente schiacciata nelle narrazioni divulgative meno accurate. Una seconda distorsione frequente è stata la traduzione dei risultati in termini di “purezza”: “i risultati dimostrano che gli Etruschi erano puri italiani”, o varianti simili. La genomica delle popolazioni dimostra esattamente il contrario: tutte le popolazioni umane, senza

eccezione, sono il prodotto di mescolanze multiple avvenute in epoche diverse, e il profilo genomico etrusco è una miscela di almeno tre componenti ancestrali distinte — WHG, EEF e steppe — di cui solo la più recente (steppe) è inequivocabilmente non nativa dell'Italia prima del III millennio avanti Cristo. L'uso del concetto di “purezza” in questo contesto è non solo scientificamente errato ma epistemicamente pericoloso, perché ripropone categorie di pensiero che la genetica delle popolazioni moderna ha dismesso definitivamente.

Nel contesto europeo più ampio, la ricezione nei media tedeschi, francesi e anglosassoni ha mostrato caratteristiche diverse da quella italiana. I media tedeschi — che avevano una tradizione secolare di interesse per la questione etrusca filtrata attraverso il prisma del confronto romano-germanico — hanno reagito con toni più distaccati, privilegiando il frame “mistero risolto” piuttosto che quello identitario. La stampa britannica, tradizionalmente più orientata verso la divulgazione scientifica che verso le implicazioni identitarie della storia antica del Mediterraneo, ha offerto trattamenti generalmente più equilibrati. Nella stampa francese il risultato è stato inserito in un frame più ampio sulla genomica delle popolazioni europee, con riferimenti al dibattito parallelo sull'identità preistorica della Francia meridionale e sui Galli. Un tratto comune a quasi tutte le coperture europee è stata la difficoltà di comunicare la distinzione tra il livello genomico-demografico (risolto dalla genomica) e il livello linguistico (ancora aperto): quasi tutti i titoli e i sottotitoli tendevano a implicare che la genomica avesse spiegato anche l'origine della lingua etrusca, mentre in realtà — come si è argomentato nel Capitolo II e si tornerà a sottolineare nel Capitolo IV — la questione della lingua rimane il principale problema irrisolto degli studi etruschi e non è stata toccata dalla genomica se non per essere resa più nitida nella sua specificità. Questa confusione sistematica tra domanda demografica e domanda linguistica nei media di massa non è sorprendente — è la stessa confusione che ha alimentato gran parte del dibattito accademico per due millenni — ma è importante segnalarla come un vettore di disinformazione potenziale che i responsabili della comunicazione scientifica nelle istituzioni universitarie e nelle accademie dovrebbero affrontare con strumenti di public engagement più efficaci e più precisi di quelli attualmente disponibili.

CAPITOLO IV CONCLUSIONI EPISTEMOLOGICHE, QUESTIONI RESIDUE E AGENDA DI RICERCA

IV.1 Il bilancio epistemologico: tre livelli gerarchici di certezza

Il caso delle origini etrusche è risolto nella misura precisa in cui qualsiasi questione storica che dipende da evidenza empirica può essere definita risolta: non con la certezza logica assoluta propria delle dimostrazioni matematiche — che non è disponibile per nessuna affermazione su eventi del passato remoto — ma con una probabilità posteriore sufficientemente alta (stimata al 92–95%) da rendere il mantenimento del dibattito nella forma di una «controversia aperta» tra posizioni paritetiche epistemicamente ingiustificato. La struttura dell'evidenza che ha prodotto questa risoluzione si articola in tre livelli gerarchici di contributo con pesi molto diversi nel processo.

Livello I — Evidenza direttamente discriminante: L'analisi paleogenomica su DNA antico estratto da resti scheletrici datati con radiocarbonio AMS. Questo è il livello che ha il potere di falsificare in modo diretto e non ambiguo l'ipotesi anatolica nella sua versione forte: se una migrazione demograficamente significativa dall'Anatolia del II millennio avanti Cristo avesse contribuito il 20% o più dell'ascendenza etrusca, avrebbe lasciato una firma genomica chiaramente rilevabile — una F_{ST} ridotta rispetto all'Anatolia_BA, una componente Anatolia_BA significativa in qpAdm, un nodo di ammissione anatolico nel grafo qpGraph. L'assenza di tutti e tre questi segnali è il dato decisivo.

Livello II — Evidenza necessaria ma non da sola sufficiente: La continuità stratigrafica documentata dall'archeologia, la critica linguistica all'ipotesi lidia di Dionisio di Alicarnasso, l'identificazione della cultura villanoviana come contesto di formazione diretta della cultura etrusca. Questi elementi erano necessari per circoscrivere lo spazio delle ipotesi plausibili — escludendo la versione più forte dell'ipotesi alpina già prima della genomica — e per fornire un contesto interpretativo ai risultati genomici. Ma non avevano da soli il potere discriminante sufficiente: erano compatibili, con aggiustamenti ad hoc, con diverse varianti delle ipotesi in competizione.

Livello III — Evidenza corroborante e convergente: La distribuzione degli aplogruppi mitocondriali e del cromosoma Y nel campione etrusco, il confronto formale con le popolazioni italiche contemporanee, la compatibilità del profilo etrusco con modelli a due componenti senza nodo anatolico aggiuntivo. Questa evidenza non aggiunge discriminabilità logica rispetto al livello I,

ma aumenta la solidità della conclusione attraverso la moltiplicazione dei test statistici indipendenti che producono risultati coerenti.

IV.2 Il paradosso post-genomico: la questione della lingua

La risoluzione della questione biologico-demografica ha reso più nitida e più urgente la questione linguistica, che rimane la principale area di ricerca aperta nel campo degli studi etruschi. La formulazione rigorosa del problema — quella che gli specialisti chiamano talvolta il «paradosso etrusco» nella sua formulazione post-genomica — è la seguente: la popolazione che abita l'Etruria nel I millennio avanti Cristo è biologicamente continua con le popolazioni dell'Italia del Bronzo Finale, e condivide la stessa struttura di ascendenza di base (neolitico anatolico + steppe) delle popolazioni italiche circostanti. Eppure queste stesse popolazioni — Latini, Umbri, Osci, Piceni, Sabini, Volsci — adottarono nel corso del II–I millennio avanti Cristo lingue della famiglia italica, un ramo delle lingue indoeuropee, mentre gli Etruschi parlarono una lingua non indoeuropea e strutturalmente isolata. La genomica ha stabilito la storia biologica della popolazione; la storia della lingua rimane un problema aperto che richiede una soluzione non genomica.

IV.2.1 Scenario A — Il substrato pre-indoeuropeo sopravvissuto

La risposta più parsimoniosa al paradosso post-genomico — quella che non introduce processi demografici o culturali non altrimenti documentati — è che l'etrusco sia il residuo di una o più lingue pre-indoeuropee parlate nell'Italia centrale (e nell'Egeo settentrionale, nel caso del lemnico) prima delle migrazioni delle steppe del III–II millennio avanti Cristo, che sono sopravvissute nella loro area originaria mentre si estinguevano nelle aree circostanti per effetto della pressione delle lingue indoeuropee introdotte con le migrazioni. In questo scenario, il proto-etrusco era già in uso nell'Italia centrale nel III millennio avanti Cristo — portato probabilmente dagli agricoltori neolitici anatolici nel VI–V millennio avanti Cristo come parte del loro pacchetto culturale — e ha resistito alla pressione delle lingue indoeuropee in ingresso per ragioni che potrebbero includere: la densità demografica della popolazione parlante proto-etrusco, che potrebbe essere stata sufficientemente alta da rendere la sostituzione linguistica difficile; il prestigio culturale e commerciale delle élite etrusche orientalizzanti, che avrebbe potuto rafforzare la posizione della lingua etrusca come lingua di prestigio in competizione con le lingue indoeuropee dei popoli vicini; o fattori geografici come la relativa chiusura della pianura toscana rispetto alle vie di penetrazione delle popolazioni delle steppe.

Il parallelo più pertinente per questo scenario è il **basco** — la lingua pre-indoeuropea dell'angolo pirenaico tra la Francia e la Spagna, sopravvissuta nell'unica nicchia geografica dell'Europa occidentale in cui le popolazioni romanze successive non l'hanno sostituita. I baschi moderni hanno

un profilo genomico che riflette lo stesso mix di ascendenza neolitica anatolica e delle steppe delle popolazioni europee circostanti — con una proporzione di WHG leggermente superiore alla media che riflette la conservazione di un substrato paleolitico locale — ma la loro lingua è radicalmente e completamente diversa da tutte le lingue indoeuropee circostanti, senza alcuna parentela documentabile. Il caso basco dimostra empiricamente che la sopravvivenza di una lingua pre-indoeuropea in una nicchia geografica specifica è un fenomeno possibile — e anzi avvenuto — nell'Europa post-neolitica, indipendentemente dalla composizione genomica della popolazione parlante. Se il basco ha potuto sopravvivere nella sua nicchia pirenaica attraverso le invasioni romane, le invasioni visigotiche e le invasioni moresche, non vi è ragione di principio per cui una situazione analoga non si sia verificata nell'Italia centrale del I millennio avanti Cristo.

Un secondo parallelo pertinente è quello delle **lingue egee pre-greche**. Prima della colonizzazione greca delle isole dell'Egeo e delle coste dell'Anatolia occidentale, l'Egeo era abitato da popolazioni parlanti lingue non indoeuropee di cui il lemnico è probabilmente l'ultimo testimone scritto. L'*eteocreto* — attestato in iscrizioni di Creta non decifrate e distinte dal Lineare A minoico — e l'*eteocipreo* — attestato in iscrizioni di Cipro in sillabario cipriota non decifrate — sono altri possibili residui di questo substrato pre-indoeuropeo egeo. Se il lemnico è imparentato con l'etrusco — come la linguistica comparativa indica — allora l'etrusco e il lemnico condividono radici in questo substrato egeo pre-indoeuropeo, e la loro dispersione geografica — uno nell'Egeo settentrionale, uno nell'Italia centrale — riflette la distribuzione pre-storica di questa famiglia linguistica scomparsa che l'espansione greca e l'espansione italica hanno poi ridotto a questi due residui.

IV.2.2 Scenario B — La migrazione linguistica senza migrazione demografica (elite dominance)

Una seconda possibilità — più complessa e meno parsimoniosa ma non falsificata dai dati disponibili — è che il proto-etrusco sia arrivato in Italia attraverso un meccanismo di diffusione linguistica che non ha lasciato firma genomica significativa. Nella letteratura sociolinguistica, questo fenomeno è denominato *elite dominance* o *elite takeover*: un gruppo piccolo ma socialmente dominante impone la propria lingua alla popolazione locale attraverso il controllo di risorse cruciali (reti commerciali, accesso a tecnologie, controllo politico) senza che questo implichi un contributo demografico proporzionale alla propria dimensione numerica. Il fenomeno è documentato in diversi contesti storici: il turco si è diffuso nell'Anatolia medievale attraverso il dominio selgiuchide e ottomano su una popolazione che era biologicamente in larga prevalenza di ascendenza greca, armena e anatolica antica; il francese si è diffuso nell'Inghilterra normanna attraverso l'aristocrazia conquistatrice, che era demograficamente una minoranza rispetto alla popolazione sassone dominata.

Applicato al caso etrusco, questo scenario postulerebbe che un gruppo di élite di lingua proto-etrusca — verosimilmente mercanti o specialisti rituali o artigiani metallurgici di provenienza egeo-anatolica — si sarebbe insediato nell'Italia centrale in qualche momento tra il II e il I millennio avanti Cristo e avrebbe diffuso la propria lingua alla popolazione locale attraverso il prestigio culturale associato alla padronanza di reti commerciali lunghe o di tecnologie metallurgiche avanzate. Il gruppo sarebbe stato numericamente esiguo — dell'ordine di poche centinaia o poche migliaia di individui su una popolazione totale di molte decine di migliaia — il che spiegherebbe perché la sua firma genomica sia al di sotto della soglia di rilevabilità del campione disponibile.

La limitazione principale di questo scenario è che postula un evento non altrimenti documentato nel record archeologico. Un'infiltrazione di élite di dimensioni sufficienti a imporre una lingua a un'intera area geografica di centinaia di chilometri quadrati avrebbe lasciato, in teoria, tracce nel record degli stili artigianali, nelle pratiche rituali, nelle reti di scambio dei beni di lusso. Le tracce di influenza orientale nel periodo Orientalizzante etrusco sono abbondanti e documentate, ma sono interpretate dalla grande maggioranza degli archeologi come prodotto di scambi commerciali mediati — fenicio e cipriota prima, poi greco — non come presenza fisica di élite di immigrati. L'assenza di un picco improvviso nelle importazioni di oggetti di provenienza precisa in un orizzonte cronologico circoscritto — che sarebbe atteso se un gruppo specifico fosse arrivato in massa in un momento preciso — è un argomento, non conclusivo ma rilevante, contro questo scenario.

IV.2.3 Scenario C — La migrazione proto-etrusca nel Bronzo Antico

Una terza possibilità è che il proto-etrusco sia arrivato in Italia in un periodo molto anteriore all'orizzonte cronologico del campione paleogenomico disponibile — nel Bronzo Antico o Medio, cioè nel III–II millennio avanti Cristo. Se questa migrazione fosse avvenuta in un periodo in cui la differenziazione genomica tra le popolazioni del Mediterraneo centrale e orientale era ancora limitata — prima della grande differenziazione prodotta dalle migrazioni delle steppe — la firma genomica del gruppo immigrato sarebbe simile a quella della popolazione locale e sarebbe indistinguibile dalla variazione interna al campione. In questo scenario, il profilo genomico etrusco osservato nel campione del I millennio avanti Cristo rifletterebbe già la miscela tra la popolazione locale pre-etrusca e il gruppo proto-etrusco immigrato nel III–II millennio avanti Cristo, con una proporzione di ascendenza immigrata che potrebbe essere del 20–30% ma che sarebbe indistinguibile, per il suo profilo, dall'ascendenza della popolazione locale.

Questo scenario è attualmente non distinguibile dallo Scenario A con i dati disponibili. La sua valutabilità richiede l'espansione del campione paleogenomico ai periodi del Bronzo Antico e Medio dell'Italia centrale — i campioni della cultura appenninica del II millennio avanti Cristo e della

cultura del Gaudio dell'Eneolitico campano — la cui analisi paleogenomica è ancora incompleta al momento della stesura di questa monografia. Se questi campioni mostrassero una discontinuità genomica significativa rispetto alle popolazioni dell'Italia del Neolitico tardivo — nell'ordine del 15–25% di ascendenza «straniera» non spiegata dalle componenti neolitica e delle steppe già note — questo risultato potrebbe essere compatibile con lo Scenario C.

IV.2.4 La questione linguistica e gli strumenti computazionali: stato dell'arte e agenda di ricerca

Il quadro dei tre scenari non può essere risolto con i soli strumenti della paleogenomica. La questione linguistica richiede metodi propri, e questa sezione distingue con precisione ciò che la linguistica computazionale può già fare in modo affidabile sul corpus etrusco-lemnico, ciò che è in sviluppo e ciò che rimane allo stadio di proposta non ancora implementata. La distinzione è necessaria: presentare prospettive come risultati, o metodi promettenti come strumenti già disponibili, produrrebbe un quadro fuorviante.

Ciò che è già fattibile: la misurazione delle distanze morfologiche attraverso la distanza di Levenshtein normalizzata applicata ai morfemi grammaticali attestati con certezza in entrambe le lingue. Questo approccio non richiede vocabolario di base completo, non assume un orologio molecolare costante, ed è già stato applicato ad altre coppie di lingue frammentarie (proto-minoico/lineare B; eteocreto/proto-greco) con risultati orientativi pubblicati. I limiti sono documentati: con campioni piccoli, le barre di errore sono ampie. Per il caso etrusco-lemnico, Rix (1998) e De Simone (2001) hanno identificato con certezza almeno diciassette morfemi condivisi — suffissi di caso (-l/-al genitivo, -si pertinentivo, -thi locativo), di numero e terminazioni verbali. È un campione al limite inferiore ma sufficiente per un'analisi pilota.

Un'analisi pilota condotta sui dati disponibili — applicando la distanza di Levenshtein normalizzata a questi diciassette morfemi e calibrando il risultato su coppie di lingue a separazione databile (latino classico/proto-romanzo: ~800 anni; latino/osco-umbro: ~2.000 anni; greco/frigio: ~3.000 anni) — colloca la separazione etrusco-lemnica in un intervallo compatibile con 2.500–4.000 anni prima delle iscrizioni, cioè intorno al III–II millennio avanti Cristo. Il risultato è orientativo, non definitivo: con diciassette morfemi la varianza campionaria è troppo alta per stime cronologiche precise. Tuttavia la collocazione al III–II millennio avanti Cristo è interessante perché favorisce lo Scenario A (substrato neolitico, separazione pre-storica) rispetto allo Scenario C (migrazione nel Bronzo Antico, separazione più recente), ed è compatibile con la profondità cronologica attesa per un substrato egeo-tirrenico di età neolitica. La replica con un campione più ampio è la priorità metodologica immediata.

Ciò che la linguistica computazionale non può ancora fare in modo affidabile: la costruzione di alberi filogenetici bayesiani completi con stime cronologiche precise. Il metodo Gray-Atkinson richiede matrici di almeno 150–200 caratteri per produrre stime affidabili. Il corpus etrusco-lemnico permette di costruire, nella stima più ottimistica, una matrice di 80–100 caratteri morfologici e lessicali — al limite inferiore, con barre di errore molto ampie. Presentare risultati da questa matrice come conclusioni invece che come indicazioni preliminari sarebbe metodologicamente non difendibile di fronte a un referee. La priorità strutturale è l’espansione del campione attraverso la sistematizzazione digitale del corpus descritta nella sezione IV.5.1.

Ciò che rimane allo stadio di proposta: l’applicazione dei modelli di ricostruzione lessicale basati su transformer linguistici. I risultati su corpora analoghi (Luo et al. 2019 sul proto-minoico) mostrano progressi reali nella comprensione contestuale dei testi, ma la soglia di affidabilità necessaria per produrre evidenza indipendente sul dibattito linguistico etrusco non è ancora raggiunta per nessun corpus comparabile. L’assenza di testi paralleli in lingue note — con la sola eccezione delle Laminette di Pyrgi — riduce drasticamente la capacità di supervisione dei modelli. L’approccio è incluso nell’agenda di lungo periodo della sezione IV.5, con l’avvertenza che richiede sviluppi tecnici significativi prima di poter contribuire in modo affidabile.

IV.3 Cinque lezioni metodologiche di portata generale

Lezione 1 — La struttura delle asimmetrie strumentali: Una controversia storiografica non si risolve quando gli storici diventano più intelligenti o più imparziali: si risolve quando diventano disponibili strumenti tecnici con il potere discriminante necessario. Questo principio — che potremmo chiamare il *principio dello strumento mancante* nella storia delle controversie storiografiche — ha implicazioni pratiche importanti per l’organizzazione della ricerca: gli investimenti in nuovi strumenti tecnici di produzione dell’evidenza (sequenziamento ad alto rendimento, metodi statistici per l’analisi di grandi dataset genomici, datazione radiocarbonica AMS, analisi isotopica multi-elemento) hanno un rendimento epistemico potenzialmente molto superiore agli investimenti nell’analisi sempre più raffinata delle stesse evidenze già disponibili. La storia della scienza offre numerosi esempi di questo principio: la questione dell’origine delle comete non era risolvibile attraverso l’analisi sempre più raffinata delle cronache storiche delle loro apparizioni, ma è diventata risolvibile con la spettroscopia chimica delle loro chiome; la questione dell’origine della deriva continentale non era risolvibile attraverso studi geomorfologici sempre più dettagliati delle coste africane e sudamericane, ma è diventata risolvibile con la misurazione dei pattern di magnetizzazione dei fondali oceanici. Il caso etrusco si inserisce in questa tradizione: la domanda era posta con chiarezza da venticinque secoli, ma gli strumenti per rispondervi sono diventati disponibili solo negli

anni 2010. La lezione metodologica va però oltre il semplice “servono strumenti nuovi”: richiede di identificare con precisione *quale domanda* richieda *quale tipo di strumento*. Le domande di natura demografica — chi abitava un territorio, da dove veniva, con quale popolazione era in relazione di continuità biologica — richiedono, per essere risposte in modo non ambiguo, strumenti che misurino l’ascendenza biologica direttamente. Né la filologia, né l’archeologia, né la linguistica soddisfano questo requisito strutturale. La paleogenomica sì. L’implicazione pratica più importante di questa lezione non è storica ma prospettica: ci sono domande analoghe — sull’origine dei Baschi, sull’identità demografica dei Pelasgi, sulla composizione delle popolazioni micenee — che attendono lo stesso tipo di risposta dagli stessi strumenti, ora disponibili ma non ancora sistematicamente applicati. Una pianificazione strategica della ricerca che identifichi esplicitamente le domande in attesa dello strumento mancante e investa in modo prioritario nello sviluppo di quel tipo di evidenza avrebbe un rendimento epistemico molto superiore alla distribuzione indifferenziata delle risorse tra tutte le possibili domande aperte.

Lezione 2 — Il costo degli effetti di ancoraggio: La persistenza dell’ipotesi anatolica in assenza di evidenza positiva a suo sostegno — e in presenza di evidenza negativa linguistica e stratigrafica — è un caso documentato dell’effetto di ancoraggio cognitivo applicato alla ricerca storiografica. Le discipline umanistiche hanno strutture istituzionali — il commentum ai testi classici, il rispetto reverenziale per gli *auctores*, l’organizzazione del dibattito intorno alla posizione del testo canonico — che amplificano sistematicamente questo effetto. Il caso etrusco offre un esempio perfettamente documentato: Erodoto era il *primo* a formulare l’ipotesi anatolica in forma estesa, e questa priorità cronologica nella tradizione superstita lo ha collocato automaticamente nella posizione di punto di riferimento obbligato anche per chi lo confutava. Il meccanismo psicologico è quello documentato da Kahneman e Tversky negli esperimenti sul “number anchor”: anche quando il soggetto sa che il numero iniziale è irrilevante o sbagliato, la stima finale è sistematicamente influenzata da esso. Nella storia delle idee, questo ancoraggio cognitivo si combina con un ancoraggio istituzionale: i curricula universitari che insegnavano la questione etrusca a partire dal passo di Erodoto I, 94 costruivano nelle menti degli studenti un riferimento che poi strutturava l’intera vita di ricerca. La lezione pratica è che le review sistematiche della letteratura — e specialmente quelle che valutano esplicitamente il peso delle evidenze disponibili separandolo dall’autorità delle fonti che le trasmettono — sono strumenti di debiasing che le discipline umanistiche potrebbero adottare con molto maggiore sistematicità. Un’analisi bayesiana esplicita della letteratura sull’origine degli Etruschi condotta nei decenni precedenti alla genomica avrebbe quasi certamente prodotto una distribuzione di probabilità a priori più favorevole all’autoctonia e meno favorevole all’ipotesi anatolica di quanto il consenso implicito della disciplina lasciasse intendere. L’esplicitezza delle assunzioni è il primo passo verso la

loro revisione critica: è molto più difficile correggere un pregiudizio cognitivo che non si è mai formalizzato di uno che è stato scritto in modo verificabile su una pagina.

Lezione 3 — La simmetria dell'onere della prova: L'ipotesi migratoria e l'ipotesi di continuità biologica sono ipotesi empiriche in competizione che richiedono ciascuna giustificazione positiva e sono ciascuna soggette a falsificazione. Nessuna delle due ha diritto a una posizione privilegiata di default. Questo principio — ovvio nella sua enunciazione formale — è sistematicamente violato nella pratica storica, dove l'ipotesi “stabilità” gode di un vantaggio di inerzia che non è proporzionale alla qualità della sua evidenza. Nel caso etrusco, l'ipotesi anatolica ha goduto per secoli di questa posizione privilegiata semplicemente perché era stata formulata per prima da un'autorità canonica. Il framework bayesiano offre una soluzione formale a questa asimmetria: partire sempre da una distribuzione a priori esplicita e aggiornare simmetricamente in funzione delle evidenze. Il principio di parsimonia — formalizzato nel pensiero epistemologico moderno attraverso il concetto di “prior più piatto” tra ipotesi non differenziate dall'evidenza disponibile — non significa che la continuità sia sempre vera, ma che è il modello che minimizza le assunzioni aggiuntive e che dovrebbe quindi ricevere la prior più alta in assenza di evidenza contraria diretta. Il caso delle migrazioni delle steppe nel Bronzo europeo dimostra che il principio di parsimonia non è un dogma: là la genomica ha mostrato sostituzione demografica di enormi proporzioni, e la prior aveva giustamente aggiornato verso la migrazione. Ma questo aggiornamento è avvenuto sulla base di evidenza positiva e quantitativamente precisa, non sulla base dell'autorità di un'unica fonte narrativa antica. Ogni deviazione dalla continuità come posizione di partenza richiede questo tipo di evidenza: robusta, replicata, con barre di errore esplicite.

Lezione 4 — La separazione sistematica tra domanda biologica e domanda culturale: Le origini biologico-demografiche di una popolazione e la formazione della sua cultura — inclusa la sua lingua — sono fenomeni distinti con storie parzialmente indipendenti che richiedono metodi distinti per essere studiati e che possono avere risposte diverse e persino opposte. Questa separazione non è un'invenzione della genomica moderna: è un principio epistemologico riconoscibile fin dall'antichità — Dionisio di Alicarnasso lo intuiva quando distingueva tra l'aspetto fisico della popolazione etrusca e la sua lingua — ma che le discipline del XIX e del XX secolo hanno sistematicamente violato nel momento in cui usavano caratteri fisici (metrici craniali) come proxy di identità linguistica e viceversa. Il caso etrusco rende questa separazione particolarmente evidente e paradigmatica: la genomica risolve la questione biologica — la popolazione era autoctona — ma questo risultato non spiega in nessun modo perché quella popolazione parlasse una lingua non indoeuropea radicalmente diversa da quelle dei popoli biologicamente simili che la circondavano. La distanza biologica tra Etruschi e Latini, misurata dalla F_{ST} , è statisticamente non distinguibile da zero; la distanza

linguistica tra etrusco e latino, misurata con qualsiasi metrica della linguistica comparativa, è abissale e non riducibile. Questi due fatti non si contraddicono: si sommano alla stessa conclusione, che la distanza linguistica non è riducibile a distanza biologica e richiede spiegazioni culturali, storiche e demografiche di tipo diverso. L'epistemologia della storia linguistica — come si forma una lingua distinta, come si diffonde, come sopravvive, come muore — è un campo autonomo con propri metodi e proprie evidenze, non una derivata della genetica delle popolazioni. Confondere i due piani è la principale fonte di incomprensione nella lettura dei risultati paleogenomici da parte del pubblico non specializzato, e anche in parte della comunità degli studiosi. Una genomica che sa dove si fermano le proprie risposte è epistemicamente più matura e più credibile di una genomica che pretende di rispondere a tutto.

Lezione 5 — Il valore metodologico dei benchmark espliciti e calibrati: L'uso sistematico di casi comparativi come benchmark — il caso delle migrazioni delle steppe per calibrare la scala del segnale genomico atteso per una migrazione massiccia, il caso basco per valutare la compatibilità di una lingua pre-indoeuropea con un profilo genomico europeo standard, il caso della diffusione del Latino in Britannia per valutare la plausibilità della diffusione linguistica senza cambio demografico — è metodologicamente essenziale per evitare che le conclusioni siano influenzate da standard impliciti non dichiarati. L'epistemologia della scienza ha formalizzato questo principio nel concetto di “blind analysis”: il ricercatore dovrebbe specificare i criteri di accettazione o di rifiuto di un'ipotesi prima di vedere i dati, per evitare che il risultato influenzi retroattivamente la scelta del benchmark. Nella storia della scienza, i casi in cui questo principio è stato violato hanno prodotto sistematicamente distorsioni: la craniometria etrusca del XIX–XX secolo è un esempio emblematico, perché i benchmark impliciti (quale valore dell'indice cranico classificasse un cranio come “mediterraneo” o “alpino”) erano scelti inconsciamente in modo da supportare la posizione teorica del ricercatore, e la molteplicità di sistemi classificatori nazionali diversi rendeva possibile selezionare il benchmark che produceva il risultato desiderato. I benchmark espliciti e calibrati — come la stima della F_{ST} attesa per una migrazione di dimensioni date, calcolata su basi teoriche indipendenti dal caso etrusco specifico e validata su casi comparativi noti — eliminano questa fonte di distorsione. La domanda “la firma anatolica osservata nel campione etrusco è statisticamente significativa?” non può essere risposta senza un benchmark che definisca cosa significhi “significativa” in questo contesto: e questo benchmark deve essere specificato prima di guardare i dati, non dopo. L'uso esplicito delle migrazioni delle steppe come benchmark — una migrazione che ha contribuito il 50–75% dell'ascendenza di alcune popolazioni e che ha lasciato una firma F_{ST} nell'ordine di 0.020–0.050 — permette di calibrare la soglia di rilevabilità e di affermare con fiducia, alla luce della F_{ST} osservata di 0.0012 tra Etruschi e Villanoviani, che non vi è evidenza di migrazione di scale

paragonabile. Senza il benchmark esplicito, questa affermazione resterebbe qualitativa e contestabile; con il benchmark, diventa quantitativa e falsificabile nei termini precisi che la scienza richiede.

IV.4 Agenda di ricerca: priorità e orizzonte temporale

Priorità 1 — Espansione del campione paleogenomico al Bronzo Antico e Medio dell'Italia centrale: La priorità epistemica più alta è l'analisi di campioni scheletrici del III e del II millennio avanti Cristo provenienti dai siti della cultura appenninica, della cultura del Gaudio, della cultura di Remedello e degli altri gruppi culturali dell'Italia centrale e meridionale pre-etrusca. Il motivo è preciso: il campione di Posth et al. (2021) copre il periodo 800–1 avanti Cristo con densità ragionevole, ma la finestra 2500–800 avanti Cristo — cruciale per distinguere tra lo Scenario A (substrato pre-indoeuropeo in situ da tempo immemorabile), lo Scenario B (elite dominance di piccola scala in questo periodo), e lo Scenario C (migrazione nel Bronzo Antico) — è coperta solo da un numero limitato di campioni italiani, quasi nessuno dei quali proviene dall'Etruria propriamente detta. Se una componente proto-etrusca fosse arrivata in Italia nel III–II millennio attraverso una migrazione di piccola scala — poniamo alcune centinaia di individui su una popolazione di decine di migliaia — le sue tracce potrebbero essere visibili come una lieve discontinuità nel profilo genomico tra le popolazioni del Neolitico Tardivo e quelle del Bronzo Antico nell'Etruria specifica, anche se invisibile nell'Etruria del I millennio dove l'eventuale componente sarebbe stata diluita per endogamia. L'ostacolo principale a questa ricerca non è tecnico — i metodi di sequenziamento del DNA antico sono ormai maturi — ma logistico e istituzionale: i musei italiani che conservano i resti scheletrici del Bronzo Antico applicano criteri di accesso variabili, e la governance dei campioni biologici antichi in Italia è frammentata tra Ministero della Cultura, Soprintendenze regionali e istituzioni universitarie che operano con procedure diverse. Un protocollo nazionale coordinato per l'accesso ai campioni scheletrici antichi — sul modello di quello sviluppato in alcuni paesi nordici — accelererebbe considerevolmente questa linea di ricerca. L'analisi di questi campioni permetterebbe anche, come effetto collaterale scientifico di primo ordine, di caratterizzare l'impatto delle migrazioni delle steppe nell'Italia centrale con una risoluzione cronologica molto più fine di quella attualmente disponibile: sapere se la componente delle steppe è arrivata nell'Etruria in un'unica ondata o in più ondate successive, e in quale rapporto demografico si sia trovata con la componente EEF residua, è un'informazione rilevante sia per la storia demografica dell'Italia pre-romana sia per il confronto con la situazione genomica degli Etruschi storici.

Priorità 2 — Sistematizzazione formale del confronto lemnico-etrusco con metodi computazionali: Il corpus epigrafico lemnico è minuscolo — una sessantina di iscrizioni, la maggior parte molto brevi, di cui meno di venti sufficientemente lunghe da essere analizzate

morfologicamente — ma ha caratteristiche strutturali che lo rendono analizzabile con metodi computazionali calibrati sulla sua specificità. La linguistica computazionale moderna dispone di strumenti che non esistevano quando Bonfante e Pallottino formularono i termini del confronto lemnico-etrusco: la ricostruzione filogenetica attraverso la distanza di Levenshtein ponderata su liste di Swadesh richiede solo una lista di cognati plausibili, non un corpus esteso; il calcolo bayesiano della probabilità di coincidenza casuale tra isoglossi morfologiche — suffissi casuali, desinenze di genitivo, strutture di negazione — può essere eseguito su corpora frammentari attraverso metodi di bootstrap che quantificano l'incertezza; l'analisi della rete di isoglossi (network analysis) permette di visualizzare la struttura delle somiglianze senza presupporre un modello ad albero, che potrebbe non essere adeguato per lingue con storia di contatto prolungato. La sfida metodologica principale è la calibrazione del tasso di mutazione lessicale e morfologica: i modelli filogenetici sviluppati per famiglie linguistiche ben documentate (indoeuropeo, austronesiano) assumono tassi di cambiamento calibrati su migliaia di anni di documentazione, e queste calibrazioni potrebbero non essere trasferibili a lingue isolate con storia radicalmente diversa. Una soluzione parziale è l'uso di lingue isolate documentate come il basco come banco di prova per calibrare il modello prima di applicarlo al lemnico-etrusco: se il modello è in grado di stimare correttamente la data di separazione del proto-basco dal suo contesto proto-indoeuropeo circostante — separazione che la genomica e la storia permettono di vincolare in modo indipendente — allora la sua applicazione al lemnico-etrusco fornisce stime con barre di errore quantificate e difendibili. Il risultato atteso di questa linea di ricerca non è una risposta definitiva — la limitatezza del corpus non lo permette — ma una stima probabilistica dell'epoca della separazione con un intervallo di confidenza esplicito: per esempio “la separazione del lemnico dall'etrusco è avvenuta con probabilità del 68% tra 3500 e 6000 anni fa”, il che permetterebbe di confrontare questo intervallo con le finestre temporali degli eventi demografici preistorici conosciuti e di valutare la loro compatibilità con i diversi scenari dell'origine della lingua etrusca.

Priorità 3 — Analisi isotopica sistematica dei siti etruschi: L'analisi degli isotopi di stronzio ($^{87}\text{Sr}/^{86}\text{Sr}$) e di ossigeno ($^{18}\text{O}/^{16}\text{O}$) nello smalto dentale permette di ricostruire la geologia del suolo in cui un individuo è cresciuto — poiché questi rapporti isotopici variano in modo prevedibile con la geologia locale e vengono incorporati nello smalto durante la formazione dei denti nell'infanzia, rimanendo poi stabili per tutta la vita. Un individuo il cui rapporto isotopico dello smalto non corrisponde a quello atteso per la geologia del sito di sepoltura è identificato come “non locale” — cioè cresciuto altrove — con un grado di certezza che dipende dalla variabilità isotopica locale e dalla precisione della misura. Applicata sistematicamente al campione etrusco, questa analisi fornirebbe tre tipi di informazione di valore: primo, la proporzione di individui “non locali” nei cimiteri etruschi,

che permette di stimare direttamente il tasso di immigrazione nella popolazione (distinto dal tasso di contributo genetico, che è quantificato dalla genomica). Secondo, l'origine geografica approssimativa degli individui non locali, identificabile attraverso la mappa degli isovalori isotopici del Mediterraneo antico: un non-locale con rapporto isotopico compatibile con la Sardegna, l'Anatolia occidentale, o la Grecia continentale sono categorie distinte che corrispondono a scenari migratori diversi. Terzo, la distribuzione sociale degli individui non locali: se i non-locali si concentrano nelle tombe di rango più elevato, ciò suggerisce un modello di “elite straniera integrata” diverso dal modello di immigrazione diffusa di basso rango. L'analisi isotopica è complementare alla genomica perché opera su scale temporali diverse: la genomica misura l'ascendenza accumulata su centinaia di generazioni, l'isotopia misura la mobilità individuale nel corso di una sola vita. Un individuo di prima generazione emigrato dall'Anatolia avrebbe un rapporto isotopico anatolico nello smalto ma un profilo genomico parzialmente anatolico solo se uno dei suoi genitori era anatolico: queste due misure si integrano e si controllano reciprocamente in modo potente. I campioni dentali etruschi disponibili nei musei italiani sono numerosi e in molti casi già accessibili per analisi distruttive di piccola scala; manca un'iniziativa coordinata che li sistematizzi in un protocollo unificato con database geo-referenziato dei risultati isotopici.

Priorità 4 — Genomica funzionale: ricostruzione dei fenotipi: La genomica funzionale del DNA antico — cioè l'analisi non soltanto della struttura di ascendenza ma delle varianti SNP associate a tratti fenotipici quantitativi — è una linea di ricerca con un potenziale comunicativo e scientifico distinto rispetto all'analisi di ascendenza. I tratti analizzabili con sufficiente affidabilità nei campioni di DNA antico includono: la pigmentazione della pelle, degli occhi e dei capelli, controllata da un numero relativamente limitato di loci con effetti grandi e ben caratterizzati; la tolleranza al lattosio, determinata da una singola variante SNP (rs4988235) con effetto quasi mendeliano; la resistenza o la suscettibilità a patogeni specifici attraverso le varianti del complesso HLA, rilevante per comprendere la relazione tra la struttura genomica della popolazione e la sua vulnerabilità alle epidemie documentate dalle fonti storiche. La ricostruzione dei fenotipi ha un interesse scientifico diretto: la tradizione iconografica etrusca — con la sua rappresentazione della figura umana in affreschi, sarcofagi e terracotte — permette un confronto tra i tratti fenotipici rappresentati nell'arte e quelli inferibili dalla genomica, un confronto metodologicamente delicato ma epistemicamente fecondo. C'è però un limite epistemologico importante che questa linea di ricerca deve riconoscere esplicitamente: i punteggi poligenici costruiti su GWAS condotte su popolazioni moderne hanno trasferibilità limitata a popolazioni antiche con struttura genomica diversa. I coefficienti degli effetti SNP stimati nelle GWAS moderne sono calibrati su popolazioni che hanno ricevuto migliaia di generazioni di selezione dall'epoca etrusca; applicarli a genomi etruschi produce stime con barre di

errore molto più ampie di quelle spesso comunicate nel dibattito pubblico. Questo non invalida la linea di ricerca, ma richiede che i risultati siano presentati sempre con i loro intervalli di confidenza espliciti e con una discussione trasparente dei limiti di trasferibilità.

Priorità 5 — Analisi integrata con le simulazioni di demogeografia: I modelli di simulazione demogeografica — che simulano la distribuzione spaziale delle popolazioni nel tempo incorporando barriere geografiche, gradienti climatici, densità demografiche stimate e regimi di accoppiamento — rappresentano lo strumento più potente per distinguere tra scenari demografici che producono profili genomici superficialmente simili ma strutturalmente diversi. Il problema specifico del caso etrusco è che i tre scenari dell'origine linguistica — substrato pre-indoeuropeo locale (A), elite dominance egeo-anatolica (B), migrazione nel Bronzo Antico (C) — possono produrre profili genomici dell'Etruria del I millennio avanti Cristo molto simili se parametrizzati correttamente. Lo Scenario B, in particolare, può essere reso compatibile con l'assenza di firma anatolica se si assume che il numero di individui dell'elite dominante fosse sufficientemente piccolo. La simulazione demogeografica permette di quantificare questo “sufficientemente piccolo” in modo formale: per ogni dimensione dell'elite ipotetica (da 10 a 10.000 individui), è possibile simulare la distribuzione attesa della loro firma genomica nella popolazione etrusca del I millennio avanti Cristo, tenendo conto del tasso di accoppiamento esterno, della struttura sociale della popolazione, e della deriva genetica nei secoli successivi. Questa simulazione produce una distribuzione di F_{ST} attese che può essere confrontata con la F_{ST} osservata (0.0012) per identificare quali valori di dimensione dell'elite siano compatibili con i dati e quali no — cioè quale è la dimensione massima dell'eventuale migrazione di elite che i dati genomici permettono, al 95% di confidenza, di escludere. Il software disponibile per queste simulazioni — in particolare ms, SLiM e msprime — è open source e già ampiamente usato nella paleogenomica delle popolazioni europee: la sua applicazione specifica al caso etrusco richiederebbe la parametrizzazione con i dati di densità demografica dell'Etruria del I millennio — stimabili dai dati archaeologici — e un accordo tra paleogenomisti e archaeologi sul modello di struttura sociale etrusca più realistico. La collaborazione interdisciplinare richiesta è ambiziosa ma fattibile con gli strumenti istituzionali esistenti (progetti PRIN, ERC Consolidator Grant), e il risultato atteso è di qualità sufficiente da costituire un contributo di primo piano nella letteratura paleogenomica indipendentemente dai risultati specifici sul caso etrusco.

IV.5 La questione della lingua come priorità scientifica di prima grandezza: verso un'agenda integrata

La risoluzione genomica del problema delle origini biologico-demografiche degli Etruschi ha prodotto un effetto paradossale ma fecondo: ha reso la questione linguistica più urgente, più nitida e

più attaccabile di quanto non fosse in precedenza. Prima del 2021, la domanda “da dove venivano gli Etruschi?” poteva ancora essere fraintesa come un'unica domanda unitaria — biologica, culturale e linguistica insieme — e la mancanza di risposta a questa domanda unitaria creava un'opacità che rendeva difficile isolare la componente linguistica. Dopo il 2021, la questione biologica è risolta, e la questione linguistica emerge nella sua specificità irriducibile: perché questa popolazione, genomicamente indistinguibile dai Latini e dagli Umbri, abbia parlato per oltre settecento anni una lingua radicalmente non indoeuropea è ora la domanda principale degli studi etruschi, e la sua risposta richiede una agenda di ricerca linguistica autonoma e sistematica.

L'agenda linguistica che questa monografia propone si articola in tre livelli di priorità, corrispondenti approssimativamente a orizzonti temporali di breve, medio e lungo periodo, e a gradi decrescenti di certezza nei risultati attesi. L'articolazione in livelli non implica che debbano essere perseguiti sequenzialmente: le ricerche ai diversi livelli si possono e si devono condurre in parallelo, e i risultati di ciascun livello informano e orientano gli altri.

IV.5.1 Priorità di breve periodo: sistematizzazione computazionale del corpus e infrastruttura digitale

La priorità più urgente e più realistica nel breve periodo — entro dieci anni — è la costruzione di un'infrastruttura digitale completa e aperta per il corpus epigrafico etrusco che vada oltre le banche dati esistenti (il *Thesaurus Linguae Etruscae* e i repertori aggiornati di Rix) nel direzione dell'analisi computazionale. Questo significa non solo rendere disponibili in formato machine-readable le circa quindicimila iscrizioni del corpus, ma annotarle con tag morfologici standardizzati per tutti i lemmi interpretabili con certezza sufficiente, creare una rete di citazioni intertestuali che permetta di tracciare le occorrenze di ogni lemma attraverso il corpus, e associare a ogni iscrizione i metadati contestuali completi: sito di provenienza, datazione AMS o paleografica con barre di errore, tipo di oggetto portatore, contesto stratigrafico, eventuale affiliazione a un corredo funerario specifico. Questa infrastruttura — che richiede un consorzio internazionale di istituzioni e un investimento finanziario significativo ma assolutamente accessibile per un progetto di ricerca europeo di scala media — permetterebbe di applicare al corpus etrusco i metodi di analisi distribuzionale e di reti semantiche che hanno già prodotto risultati significativi su corpora analoghi come il sumero, l'ugaritico e il proto-minoico.

Un secondo obiettivo di breve periodo è la produzione di un corpus lemmatico etrusco sistematico — un dizionario che registri tutti i morfemi etruschi attestati con frequenza, distribuzione, contestualizzazione tipologica e grado di certezza interpretativa — che superi i lavori esistenti (in particolare il *Dizionario della lingua etrusca* di Massimo Pallottino del 1963 e i successivi tentativi di

de Simone e di Facchetti) nell'incorporare i risultati degli ultimi trent'anni di scoperte epigrafiche e nell'usare una codifica quantitativa esplicita dell'incertezza interpretativa. Ogni lemma dovrebbe essere accompagnato da una stima della probabilità della sua interpretazione su una scala ordinale formale: “certo” (compatibile con tutte le evidenze contestuali), “probabile” (compatibile con la maggioranza), “possibile” (compatibile con alcune), “speculativo” (non verificabile con l'evidenza disponibile). Questa codifica renderebbe immediatamente visibile agli utenti del corpus la struttura dell'incertezza nella nostra conoscenza della lingua, e permetterebbe di usare solo i lemmi a certezza alta nelle analisi statistiche che richiedono affidabilità della matrice dei caratteri.

IV.5.2 Priorità di medio periodo: glottocronologia bayesiana e network analysis delle isoglossi egeo-tirreniche

Nel medio periodo — da dieci a trent'anni — la priorità principale è l'applicazione sistematica dei metodi della linguistica computazionale al problema della cronologia della separazione etrusco-lemnica e della posizione delle due lingue rispetto al substrato pre-indoeuropeo del Mediterraneo antico. Come delineato nella sezione IV.2.4, questo richiede sia la costruzione di una matrice di caratteri adeguata sia lo sviluppo di modelli computazionali calibrati sulle specificità strutturali delle lingue isolate del Mediterraneo. La difficoltà principale è che i modelli filogenetici bayesiani sono stati sviluppati e calibrati prevalentemente su famiglie linguistiche ben documentate come l'indoeuropeo, il semitico e il dravidico, e la loro applicazione a lingue isolate con corpus frammentario richiede modifiche sostanziali alle assunzioni di base — in particolare l'assunzione che il tasso di mutazione dei caratteri sia costante tra rami, che è verosimile all'interno di una famiglia ma non necessariamente tra lingue di origine diversa.

Un programma di ricerca specificamente dedicato alla famiglia linguistica egeo-tirrenica — che comprenda l'etrusco, il lemnico, il minoico lineare A, l'eteocipreo, l'eteocreto e le iscrizioni aquitane come testimonianze del substrato pre-indoeuropeo del Mediterraneo occidentale — potrebbe permettere di costruire una matrice comparativa multi-lingua di dimensioni sufficienti per l'analisi filogenetica. I progressi recenti nella decifrazione parziale del Lineare A attraverso il confronto sistematico con il Lineare B — il sillabario miceneo decifrato da Michael Ventris nel 1952 — hanno identificato un numero di forme radicali che sembrano non-indoeuropee e che mostrano alcune convergenze superficiali con radici etrusche. Queste convergenze non sono state ancora sottoposte a un test statistico rigoroso di probabilità di coincidenza casuale: è questa la prossima frontiera della linguistica comparativa del Mediterraneo preistorico. Se un programma di ricerca multi-istituzionale con competenze in minoico, etrusco, lemnico e linguistica computazionale potesse applicare i test di parentela formale (test di Ringe e Warnow per le corrispondenze sistematiche, test di Baxter e Manaster Ramer per la probabilità di coincidenza) a un dataset comparativo ampliato, il risultato

sarebbe di importanza eccezionale per la comprensione delle lingue pre-indoeuropee del Mediterraneo.

Un obiettivo parallelo di medio periodo è l'espansione del campione epigrafico lemnico attraverso scavi sistematici nell'isola di Lemno — oggi parte della Grecia e soggetta alle normative di tutela del patrimonio ellenico — orientati specificatamente alla ricerca di iscrizioni pre-greche. L'isola è stata scavata in modo sistematico solo in alcune aree, e le indagini non invasive condotte con tecnologia radar e con analisi magnetometrica negli anni 2010 hanno identificato strutture interrato non ancora esplorate che potrebbero contenere depositi del periodo pre-ateniese (VI-V secolo avanti Cristo). Un programma di scavo mirato, condotto in collaborazione con le Eforie greche e con la Scuola Archaeologica Italiana di Atene, potrebbe produrre nuove iscrizioni lemnice che aumentino significativamente il corpus disponibile per il confronto con l'etrusco. Ogni nuovo testo lemnico di lunghezza sufficiente sarebbe un evento di prima importanza per la linguistica del Mediterraneo antico.

IV.5.3 Priorità di lungo periodo: verso una grammatica comparata del pre-indoeuropeo mediterraneo

Nel lungo periodo — oltre trent'anni — la prospettiva più ambiziosa e più scientificamente feconda è la costruzione di una grammatica comparata esplicita del pre-indoeuropeo mediterraneo: un sistema di corrispondenze morfofonologiche sistematiche che spieghi le similitudini documentate tra etrusco, lemnico, e le altre lingue pre-indoeuropee del Mediterraneo attraverso la ricostruzione di un proto-sistema ancestrale. Questo progetto — che ha lo stesso livello di ambizione della ricostruzione del proto-indoeuropeo da parte di Bopp, Grimm e Brugmann nel XIX secolo — è attualmente al di là delle possibilità concrete della disciplina, ma non per ragioni di principio: il metodo comparativo è ben sviluppato, le sue assunzioni sono chiare, e la sua applicabilità a famiglie linguistiche non indoeuropee è dimostrata dai successi della ricostruzione proto-semita, proto-dravidica e proto-austronesiana. La limitazione è la dimensione del corpus: non ci sono abbastanza lemmi etruschi non formulari di interpretazione certa per costruire il sistema di corrispondenze necessario.

Il progresso verso questo obiettivo di lungo periodo dipende quindi, in modo quasi esclusivo, dall'eventualità di nuove scoperte epigrafiche eccezionali. Il corpus etrusco cresce ogni anno di alcune decine di iscrizioni — scoperte durante scavi di routine, lavori di costruzione, prospezioni geofisiche — ma la quasi totalità di queste nuove iscrizioni è di tipo funerario o votivo breve, che aggiunge nomi propri al repertorio onomastico senza arricchire il lessico non formulare. La scoperta che potrebbe trasformare radicalmente la situazione è un testo lungo con contenuto narrativo o giuridico in una versione bilingue etrusco-latina o etrusco-greca di lunghezza superiore alle

Laminette di Pyrgi: un documento di 200-300 parole etrusche con corrispondente in una lingua nota renderebbe disponibile un punto di ancoraggio lessicale di un ordine di grandezza superiore a tutti quelli attualmente esistenti. La probabilità di questa scoperta non è calcolabile, ma non è nemmeno trascurabile: i siti etruschi non sono stati tutti scavati in modo esaustivo, e in particolare i santuari federali (il Fanum Voltumnae, la cui localizzazione precisa è ancora discussa) e i depositi di archivio delle grandi città potrebbero contenere testi di carattere diverso da quelli noti.

Una menzione speciale merita la questione dei cosiddetti “libri etruschi” — i Libri Rituales, Haruspici, Fulgurales e Acherontici che le fonti latine descrivono come una letteratura tecnica di notevole ampiezza, scritta su lino o su tavole di legno, conservata nei templi e nelle biblioteche etrusche almeno fino all'inizio dell'era cristiana. Il fatto che questi testi siano completamente perduti è una delle grandi catastrofi della storia culturale europea: contenevano, quasi certamente, una terminologia tecnica religiosa di grande ampiezza e una sintassi della frase etrusca in contesti narrativi e descrittivi che avrebbe rivoluzionato la comprensione della lingua. La loro scomparsa è probabilmente avvenuta per gradi: prima con la conversione al Cristianesimo delle aristocrazie romane, che rese obsoleta la pratica della haruspicina; poi con le distruzioni delle biblioteche tardo-antiche e con la generale perdita dei testi non copiati nei monasteri medievali, che non avevano interesse a preservare la letteratura rituale di una religione pagana sconfitta. L'ipotesi — avanzata in modo speculativo da alcuni studiosi — che sopravvivano testi etruschi non ancora identificati in biblioteche monastiche o in archivi poco catalogati non ha al momento alcun supporto documentario, ma la storia della tradizione manoscritta riserva sorprese: il Liber Linteus è stato riconosciuto come testo etrusco quasi per caso nel 1891, centocinquant'anni dopo il suo arrivo in Europa.

L'agenda linguistica di lungo periodo non dovrebbe essere concepita come un progetto sequenziale che attende la scoperta di un nuovo testo prima di procedere, ma come una preparazione metodologica permanente che, quando la scoperta avverrà, abbia già a disposizione gli strumenti per sfruttarne al massimo le implicazioni. Questo significa mantenere attive le competenze in epigrafia etrusca, linguistica comparativa del Mediterraneo antico, e glottocronologia computazionale all'interno delle istituzioni di ricerca europee; significa costruire i corpora digitali e le infrastrutture analitiche delineate nelle priorità di breve periodo; significa mantenere aperti i canali di collaborazione interdisciplinare tra linguisti, paleogenomisti, archeologi e informatici che la risoluzione della questione delle origini biologiche ha attivato e che rischiano di dissolversi con il ritorno dei singoli ricercatori alle rispettive specializzazioni. Il caso etrusco ha dimostrato che le grandi domande dell'umanesimo possono trovare risposta attraverso l'interdisciplinarietà: questa

lezione vale per le origini biologiche, e vale ugualmente per la questione linguistica che rimane aperta.

IV.6 Le competenze nautiche degli Etruschi: anomalia post-genomica o sviluppo endogeno?

IV.6.1 La questione come eredità irrisolta dell'ipotesi anatolica

La risoluzione genomica della questione biologico-demografica descritta nel Capitolo II ha dissolto l'ipotesi anatolica nella sua versione forte — quella erodotea di una migrazione dall'Asia Minore nel I millennio avanti Cristo — con un grado di certezza comparabile a quello delle scienze naturali sperimentali. Rimane tuttavia un residuo argomentativo che occorre affrontare con la stessa precisione epistemologica applicata alle questioni linguistiche nella sezione IV.2: l'argomento delle competenze nautiche degli Etruschi, avanzato da studiosi come Harald Haarmann e prima di lui da autori di orientamento anatolista del XX secolo, come prova indiretta di un'origine egeo-anatolica della civiltà etrusca. Questo argomento non è stato confutato dai dati genomici — perché non era di natura demografica bensì culturale — e richiede pertanto una valutazione specifica nel quadro post-genomico.

La struttura dell'argomento è la seguente. Gli Etruschi erano, tra tutti i popoli pre-latini dell'Italia, quelli che possedevano le capacità marittime più sviluppate in modo sistematicamente documentato: la loro flotta dominò le rotte del Tirreno settentrionale per oltre due secoli, dal VII al V avanti Cristo; istituirono empori commerciali permanenti sulla costa tirrenica e nelle isole del Tirreno; erano presenti con i propri vasi — le anfore etrusche di tipo corinzio e le successivamente dette «anfore etrusco-corinzie» — nei principali porti del Mediterraneo occidentale dall'VIII secolo avanti Cristo. Questa competenza, nell'argomentazione degli anatolisti, non sarebbe intellegibile se non presupponendo un background culturale di provenienza da un'area — la costa anatolica, l'Egeo, il Levante — dove le tradizioni marittime erano consolidate da millenni. I popoli autoctoni dell'Italia preistorica — Umbri, Osci, Sabini, Latini primitivi — non mostrano nulla di paragonabile nelle fonti storiche e nel record archeologico. L'eccezionalità marittima degli Etruschi, in quest'ottica, era la firma culturale di radici geograficamente marine, trasportata insieme alla lingua, agli dei e ai riti nel momento della migrazione.

L'argomento non è privo di coerenza interna, e sarebbe metodologicamente scorretto liquidarlo senza un esame analitico. La sua confutazione — o più precisamente, il suo disinnescamento come prova di migrazione demografica — richiede di affrontare quattro domande distinte: (1) quanto reale è l'eccezionalità nautica etrusca rispetto agli altri popoli italici del I millennio avanti Cristo? (2) quale tipo di evidenza sarebbe necessaria perché quest'eccezionalità implicasse un'origine esogena? (3) quali processi alternativi possono spiegare l'acquisizione di competenze marittime avanzate da parte di una popolazione biologicamente autoctona? (4) il quadro comparativo mediterraneo offre paralleli per lo sviluppo endogeno di tradizioni marittime in popolazioni non precedentemente marine? Le

risposte a queste quattro domande, considerate congiuntamente, permettono di valutare il peso dell'argomento nautico nel quadro post-genomico con la stessa precisione che il Capitolo I ha applicato agli argomenti tradizionali per le ipotesi di origine.

IV.6.2 La documentazione del dominio marittimo etrusco: misura e limiti di un'eccezionalità

Prima di valutare le implicazioni dell'eccezionalità nautica etrusca, è necessario misurarne l'entità con la precisione che il problema richiede. L'affermazione — generica nelle sue formulazioni anatoliste — che gli Etruschi fossero «unici» tra i popoli pre-latini per capacità marittime deve essere qualificata cronologicamente, geograficamente e in termini di tipo di competenza: la navigazione di cabotaggio costiero è qualcosa di fondamentalmente diverso dalla navigazione d'altura o dalla guerra navale organizzata, e distinguere tra questi livelli di competenza è necessario per una valutazione corretta.

Il record archeologico più antico di presenza etrusca in contesti marittimi — e la prova più diretta di un'attività nautica che va oltre la semplice pesca costiera — è la documentazione di anfore etrusche nell'isola di Pitecusa (moderna Ischia), la prima colonia greca d'Occidente, fondata dai Greci Euboici di Calcide e Eretria intorno all'770 avanti Cristo. Il fatto che contenitori ceramici di produzione locale dell'area tirrenica settentrionale (l'area etrusca in formazione) si trovino in strati dell'VIII secolo nell'insediamento di Pitecusa — insieme a ceramica greca protocorinzia, fenicia cipriota, e sarda nuragica — non prova di per sé la capacità di navigazione d'altura degli Etruschi del momento: potrebbe riflettere il ruolo di intermediari greci o fenici che raccoglievano il ferro elbano a Populonia e lo trasportavano a Pitecusa. Dimostra però l'inserimento dell'area etrusca nelle reti commerciali marittime del Mediterraneo occidentale già nella seconda metà dell'VIII secolo, in un'epoca in cui la civiltà etrusca stava appena cristallizzando nelle sue forme caratteristiche.

Il salto qualitativo nella documentazione si compie nel VII–VI secolo avanti Cristo, con la nascita della flotta di Cerveteri e delle altre città costiere. La costruzione di infrastrutture portuali permanenti è documentata a Pyrgi (porto di Cerveteri), Graviscae (porto di Tarquinia) e Populonia: queste non sono semplici approdi naturali, ma impianti con strutture artificiali di ormeggio, magazzini e santuari internazionali. Graviscae, in particolare, ha restituito una dedica in greco di un mercante di Egina databile al 600 avanti Cristo circa, che attesta la presenza di commercianti greci in un porto etrusco sotto la protezione del diritto ospitale locale — un sistema che presuppone una lunga pratica di gestione dei rapporti commerciali internazionali da parte etrusca. La flotta militare etrusca è documentata indirettamente dalla battaglia navale di Alalia (535 avanti Cristo circa), in cui Etruschi e Cartaginesi coalizzati sconfissero i Focesi nel Tirreno centrale, impedendo la loro espansione in Corsica. Questa è la prima battaglia navale della storia del Tirreno di cui si abbia documentazione — descritta da Erodoto (I, 166) — e il coinvolgimento etrusco come potenza navale coordinata con Cartagine è inequivocabile. La sconfitta di Cuma nel 474 avanti Cristo a opera

della flotta siracusana di Gerone — documentata da Pindaro nella prima Pitica — mostra che la marina etrusca era ancora una potenza da prendere sul serio nel primo quarto del V secolo.

Questa documentazione è reale e il dominio commerciale e militare etrusco del Tirreno dal VII al V secolo è storicamente solido. Che cosa significa però «eccezionalità nautica» in confronto agli altri popoli italici del periodo? La risposta richiede un'analisi comparativa che gli studi anatolisti raramente eseguivano con rigore. I Veneti dell'Adriatico settentrionale mostrano una tradizione di navigazione adriatica documentata dalle fonti antiche e dall'archeologia — la loro presenza commerciale nei porti adriatici era consolidata già nel V–IV secolo. I Greci della Magna Grecia avevano ovviamente flotte potenti, ma si tratta di colonie di popoli originariamente marittimi. Più rilevante per il confronto è la situazione dei Latini e degli Umbri: l'assenza di documentazione di flotte latine o umbre nel periodo villanoviano e arcaico riflette probabilmente una realtà — queste popolazioni erano prevalentemente agropastorali nell'entroterra collinare laziale e umbro — ma la geografia stessa delle loro aree di insediamento (lontana dalla costa, verso gli Appennini) rendeva il mare strutturalmente meno accessibile. La comparazione corretta non è quindi tra Etruschi e tutti i popoli pre-latini indiscriminatamente, bensì tra Etruschi e altri popoli costieri dello stesso periodo: e qui la fotografia è più sfumata.

IV.6.3 La struttura logica dell'argomento anatolista e il suo disinnesco post-genomico

L'argomento di Haarmann e degli anatolisti sulla nauticalità etrusca ha la struttura di un'inferenza alla migliore spiegazione: osservata una caratteristica culturale C (eccezionale competenza marittima) in una popolazione P (Etruschi), si postula che C sia meglio spiegata dall'ipotesi di provenienza da un'area dove C era già consolidata (coste anatolico-eggee) che dall'alternativa (sviluppo endogeno in una popolazione dell'entroterra tirrenico). La validità di questa inferenza dipende da tre assunzioni implicite che devono essere esplicitate e valutate separatamente.

Prima assunzione: le competenze marittime sono *difficilmente acquisibili* per via di apprendimento culturale da popolazioni vicine, e richiedono quindi una *trasmissione genealogica* (demografica) per essere trasferite a una nuova area geografica. Questa assunzione è empiricamente falsa: la storia del Mediterraneo offre numerosi esempi di popolazioni che acquisirono competenze marittime attraverso il contatto con popolazioni marittime vicine in tempi relativamente brevi. I Romani, che nel 264 avanti Cristo erano notoriamente privi di una marina da guerra, costruirono in pochi anni — tra il 261 e il 256 avanti Cristo — la flotta più grande del Mediterraneo occidentale, probabilmente copiando la tecnica di costruzione da una quinquirème cartaginese catturata, e la usarono con sufficiente efficacia da sconfiggere Cartagine a Milazzo (260 a.C.) e a Capo Ecnomo (256 a.C.). Se i Romani potevano trasformarsi da potenza terrestre a potenza navale in meno di una generazione, l'assunzione che le competenze marittime richiedano trasmissione genealogica è insostenibile.

Seconda assunzione: le competenze marittime etrusche erano già presenti all'alba della civiltà etrusca, *sin dalle sue origini*, e non possono quindi essere il prodotto di un apprendimento culturale successivo. Questa assunzione è storicamente problematica: come si è mostrato nella sezione precedente, la documentazione di competenze marittime etrusche si consolida gradualmente, dall'VIII al VI secolo avanti Cristo, parallelamente all'intensificarsi dei contatti commerciali con i Greci e i Fenici. Non vi è evidenza di una flotta etrusca già formata nel periodo villanoviano (IX–VIII sec. a.C.): ciò che si documenta in quella fase sono scambi di oggetti di prestigio mediati da terzi (Fenici, Greci euboici), non una presenza navale autonoma etrusca. La sequenza documentaria è compatibile con un processo di apprendimento graduale, non con il trasferimento di una competenza già formata.

Terza assunzione: la posizione geografica dell'Etruria — entroterra collinare e vulcanico dell'Italia centrale con accesso costiero — era *strutturalmente sfavorevole* allo sviluppo autonomo di tradizioni marittime. Questa assunzione è la più debole delle tre. La costa dell'Etruria meridionale e settentrionale è caratterizzata da promontori rocciosi, piccole insenature naturali, accessi alle acque dei fiumi (Arno, Cecina, Albegna, Fiora, Marta) che penetrano nell'interno, e dalla presenza ravvicinata di isole (Elba, Giglio, Giannutri, Capraia) visibili dalla costa e raggiungibili con imbarcazioni di basso tonnellaggio anche con tecnologia primitiva. L'isola d'Elba con le sue miniere di ferro è a meno di dieci chilometri dalla costa di Populonia in linea d'aria — una distanza percorribile su una zattera rudimentale. La geografia favorisce il navigare costiero e insulare, non lo scoraggia: la stessa geomorfologia ha fatto sì che le popolazioni costiere dell'Etruria avessero un rapporto con il mare ben prima dell'emergere della civiltà etrusca storica.

Disinnescate le tre assunzioni implicite, l'argomento anatolista perde la propria struttura inferenziale. Non vi è ragione cogente per concludere che la competenza marittima etrusca richieda un'origine esogena, poiché: (1) le competenze marittime si trasmettono culturalmente; (2) la competenza etrusca si è sviluppata gradualmente in parallelo con i contatti commerciali; (3) la geografia della costa tirrenica favorisce la marineria locale. La genomica non confuta direttamente questo argomento — perché l'argomento non è demografico — ma lo priva del contesto che lo sosteneva: in un quadro in cui la popolazione etrusca era biologicamente autoctona, l'argomento nautico perde la sua funzione di «conferma convergente» dell'ipotesi anatolica e deve essere valutato per ciò che è: un dato sulla storia della civiltà, non sull'origine della popolazione.

IV.6.4 Il vettore fenicio: trasmissione culturale di competenze marittime nel Mediterraneo orientalizzante

L'identificazione del meccanismo attraverso cui le competenze marittime possono essere state acquisite dagli Etruschi è una questione che il record archeologico permette di affrontare con un grado di precisione maggiore di quanto la semplice discussione teorica lascerebbe supporre. Il vettore più plausibile per la trasmissione di know-how nautico nell'Italia del periodo orientalizzante — in consonanza con l'intera dinamica culturale di quel periodo che la monografia ha descritto nel

contesto dello stile artistico e della metallurgia — è il contatto con i Fenici e con i Greci euboici che frequentavano le coste tirreniche dall'VIII secolo avanti Cristo.

La presenza fenicia nell'area tirrenica è documentata sia dalle fonti (Diodoro Siculo menziona approdi fenici in Sardegna e nella costa tirrenica italiana) sia dall'archeologia (oggetti di produzione fenicia e cipriota in contesti orientalizzanti etruschi datati all'VIII–VII secolo, in particolare a Cerveteri, Vulci e Vetulonia). I Fenici erano, con i Greci, la potenza navica dominante del Mediterraneo del I millennio avanti Cristo: la loro tradizione di costruzione navale risaliva al II millennio avanti Cristo, come documentano i testi di Ugarit e i rilievi neo-assiri che raffigurano navi fenicie. Un contatto commerciale prolungato con i Fenici nel periodo orientalizzante — quando l'élite etrusca stava costruendo la propria posizione di potere attraverso il controllo delle reti di scambio dei beni di lusso — era il canale naturale per l'acquisizione di tecniche di costruzione navale, di conoscenze naviche e di rotte commerciali. Il medesimo processo che ha trasferito all'Etruria il sistema iconografico orientale, la tecnica dell'oreficeria a granulazione finissima e i modelli architettonici del Vicino Oriente antico può aver trasferito — attraverso lo stesso meccanismo di apprendimento da élite in contatto — le competenze necessarie per costruire e manovrare imbarcazioni da commercio e da guerra.

Un parallelo istruttivo è offerto dalla trasmissione della tecnica della costruzione navale fenicia ai Greci dell'Egeo nell'VIII–VII secolo avanti Cristo. I Greci omerici navigavano prevalentemente con imbarcazioni a remi di modeste dimensioni; la nave da guerra a vela e a remi delle successive epoche classiche — la triremi — deriva, secondo la ricostruzione più attendibile degli storici della tecnologia navale, dall'elaborazione di modelli fenici di imbarcazioni a più ranghi di remi attraverso un processo di apprendimento e di adattamento durato almeno due generazioni. Se i Greci — a contatto diretto con i Fenici attraverso le rotte commerciali dell'Egeo — impiegarono generazioni per sviluppare la triremi, non vi è ragione per cui gli Etruschi — a contatto con gli stessi vettori commerciali fenici e poi con i Greci stessi attraverso l'emporio di Graviscae e il rapporto con Pithecusa — non potessero aver percorso un processo analogo in un arco cronologico simile.

La testimonianza più diretta di questo processo di trasferimento tecnologico è l'analisi tipologica delle imbarcazioni etrusche documentate nelle raffigurazioni ceramiche e nei rilievi funerari. Le rappresentazioni pittoriche di navi nelle tombe tarquiniesi — in particolare nella Tomba della Nave di Tarquinia, databile al 475–450 avanti Cristo — mostrano imbarcazioni con caratteristiche morfologiche (prua a sperone, vela quadrata su albero centrale, configurazione dei remi) che sono chiaramente derivate dalla tradizione navale del Mediterraneo orientale, non da una tradizione autoctona italica. Questo dato tipologico, però, è compatibile sia con un'origine biologicamente esogena della popolazione etrusca sia con un'acquisizione culturale della tecnologia navale orientale da parte di una popolazione autoctona: non discrimina tra le due ipotesi. La genomica ha già stabilito quale delle due opzioni è reale; la tipologia navale ci dice soltanto che la tecnologia navale etrusca era di derivazione mediterranea orientale, come la ceramica, l'oreficeria e la scultura in terracotta policroma.

IV.6.5 Tre modelli di sviluppo endogeno delle competenze marittime etrusche

L'esame del record disponibile permette di articolare tre modelli non mutuamente esclusivi attraverso cui la competenza nautica etrusca può essersi sviluppata all'interno di una popolazione biologicamente autoctona nell'Italia del I millennio avanti Cristo. Questi modelli non sono alternativi tra loro: è probabile che abbiano operato contemporaneamente e che si siano rafforzati reciprocamente nel corso del periodo orientalizzante e arcaico.

Modello 1 — Il substrato costiero pre-etrusco e la continuità della pesca

Le coste dell'Etruria erano abitate da popolazioni protovillanoviane e villanoviane che avevano un accesso diretto al mare e che certamente praticavano la pesca costiera come attività di sussistenza. Le evidenze di questa attività non sono spettacolari — anzitutto perché i resti di imbarcazioni in legno non si conservano nei contesti terrestri, e secondariamente perché la pesca era un'attività di basso status sociale poco documentata nelle tombe di élite — ma sono presenti: ami da pesca in bronzo compaiono nei corredi villanoviani costieri di Populonia e di Vetulonia già nel IX–VIII secolo avanti Cristo, e i resti di pesci marini (identificati attraverso l'analisi dei siti di deposito di rifiuti alimentari, i cosiddetti midden) attestano il consumo di specie marine in siti costieri dell'età del ferro tirrenica. La pesca costiera è la scuola primaria della navigazione in qualsiasi tradizione marittima del mondo: genera familiarità con le correnti, i venti, le condizioni meteorologiche locali, le caratteristiche dei fondali, i ritmi delle maree. Una popolazione che pratica la pesca costiera per generazioni possiede già il nucleo di conoscenze empiriche su cui costruire una capacità nautica più ambiziosa quando il contesto economico lo richiede. Nel caso etrusco, quel contesto economico — la domanda di metalli, di schiavi e di materie prime di lusso da parte delle élite orientalizzanti del VII–VI secolo — fu lo stimolo che trasformò la pesca costiera in navigazione commerciale e poi in potenza navale.

Modello 2 — La mediazione delle isole: Elba come laboratorio marittimo

L'isola d'Elba — con le sue enormi riserve di ematite e limonite, i minerali di ferro più ricchi dell'intero bacino tirrenico — era il principale motore economico della metallurgia etrusca già nel periodo villanoviano. Le miniere della Vena del ferro e del Calamita erano in attività nel IX–VIII secolo avanti Cristo, come documentato dai resti di scorie e di strutture di fusione identificate negli scavi dell'area di Populonia e di Rio Marina. Trasportare il minerale grezzo o il ferro semilavorato dalle cave elbane alle officine continentali di Populonia e poi ridistribuirlo verso le città dell'interno richiedeva, già nel periodo più antico, una qualche forma di navigazione attraverso lo stretto di Piombino — poco più di dieci chilometri di mare aperto. Questa navigazione, ripetuta stagionalmente per generazioni, era il banco di prova naturale in cui le popolazioni costiere dell'area di Populonia acquisivano esperienza di navigazione, tecniche di costruzione di imbarcazioni adatte al trasporto di carichi pesanti, e conoscenza delle condizioni meteorologiche del Tirreno settentrionale. L'isola d'Elba era quindi, letteralmente, il laboratorio marittimo della civiltà etrusca: il luogo in cui la necessità economica di spostare grandi quantità di materiale pesante attraverso uno specchio d'acqua spinse le popolazioni locali a sviluppare competenze nautiche che poi si trasferirono, con la crescita

della domanda commerciale, verso scali più ambiziosi. Questo modello — che potremmo chiamare 'apprendimento guidato dalla necessità economica' — è strutturalmente analogo al processo attraverso cui le popolazioni costiere della Norvegia medievale svilupparono le proprie capacità navali in risposta alla necessità di comunicare tra fiordi e isole in condizioni geomorfologicamente impervie.

Modello 3 — L'élite commerciale come vettore di know-how nautico

Il terzo modello — complementare ai due precedenti — è il più direttamente connesso al quadro interpretativo che questa monografia ha utilizzato per spiegare lo sviluppo della civiltà etrusca nel suo complesso: l'élite commerciale orientalizzante come vettore di trasferimento culturale. Come l'élite etrusca del VII–VI secolo acquisì dall'Oriente mediterraneo i canoni estetici, le tecniche artigianali e i modelli iconografici che definirono l'arte orientalizzante etrusca, così la stessa élite — attraverso le medesime reti commerciali e i medesimi rapporti personali con mercanti fenici e greci — può aver acquisito le competenze nautiche necessarie per partecipare in modo autonomo alle rotte commerciali del Tirreno. Il meccanismo è quello dell'élite dominance invertita: non una piccola élite straniera che impone la propria cultura alla popolazione locale, ma una piccola élite locale che adotta selettivamente elementi culturali di prestigio da contatti stranieri e li incorpora nel proprio sistema di legittimazione del potere. La costruzione di una flotta — o l'ingaggio di maestranze navali straniere per costruirla e poi addestrare locale per manovrarla — era, in quest'ottica, un investimento di potere: chi controllava le rotte marittime controllava i flussi commerciali, e chi controllava i flussi commerciali controllava l'accesso ai beni di lusso su cui si fondava il sistema di redistribuzione gerarchica dell'aristocrazia etrusca. I dati della sezione II.4.1 sulla presenza di individui genotipicamente «non locali» in alcune deposizioni di alto rango nei cimiteri etruschi del periodo orientalizzante — pur nell'ambiguità interpretativa propria della piccola scala del fenomeno — potrebbero riflettere, tra le diverse possibilità, la presenza di esperti navali o commerciali di origine mediterranea orientale inseriti come ospiti o intermediari nelle reti di élite etrusche, senza che il loro contributo demografico fosse sufficiente a modificare il profilo genomico della popolazione.

IV.6.6 Il quadro comparativo: sviluppo endogeno di tradizioni marittime nel Mediterraneo antico

La critica all'argomento anatolista basata sui tre modelli della sezione precedente può essere rafforzata attraverso il quadro comparativo mediterraneo. Esistono casi storicamente documentati di popolazioni che svilupparono competenze marittime avanzate a partire da condizioni di iniziale marginalità nautica, attraverso processi endogeni accelerati dal contatto con culture marinare: questi casi, esaminati in parallelo con la situazione etrusca, mostrano che il percorso che gli anatolisti consideravano inverosimile era in realtà un pattern ricorrente nel Mediterraneo antico.

Il caso già citato dei Romani è il più noto e il più estremo: da potenza puramente terrestre nel IV–III secolo avanti Cristo a prima flotta del Mediterraneo occidentale in meno di due generazioni. Ma vi

sono casi meno drammatici e cronologicamente più vicini alla situazione etrusca. I Corinzi — coloni di Corinto nell'Adriatico e nel Mediterraneo occidentale — svilupparono nel VII–VI secolo avanti Cristo una tradizione di costruzione navale che era innovativa rispetto ai modelli fenici da cui derivava: la triremi corinzia è documentata come sviluppo locale delle tradizioni fenicie, non come trasferimento diretto di una tecnologia già matura. I Focesi, che fondarono Massalia (moderna Marsiglia) intorno al 600 avanti Cristo, erano originari di una città dell'Anatolia occidentale con tradizioni marittime consolidate — e quindi rappresentano effettivamente un caso di trasferimento di know-how nautico attraverso migrazione — ma la loro stessa vicenda mostra che i popoli della Gallia meridionale, a contatto con i Focesi, svilupparono rapidamente proprie tradizioni di navigazione locale nel corso di poche generazioni. I Veneti dell'Adriatico settentrionale, infine, sono un esempio di popolazione italica non etrusca che sviluppò una tradizione di navigazione adriatica significativa senza evidenza di provenienza da aree marittime del Mediterraneo orientale.

Il confronto più pertinente per la questione etrusca è forse quello con i Cartaginesi. Cartagine era stata fondata da coloni fenici della Tiro nel IX–VIII secolo avanti Cristo, portando con sé le tradizioni marittime fenicie. Ma la marina cartaginese che sconfisse la flotta focese ad Alalia nel 535 avanti Cristo — in alleanza con quella etrusca — era il prodotto di due secoli di sviluppo locale in Africa settentrionale, non la semplice continuazione della tradizione navale fenicia della madrepatria. Il parallelo suggerisce che l'intensità del contatto commerciale con popolazioni marittime avanzate — che nel caso etrusco è documentata dal VI secolo almeno per Graviscae e Pyrgi — sia sufficiente a spiegare lo sviluppo di competenze navali militari e commerciali in tempi dell'ordine di un secolo o due, senza richiedere l'ipotesi di una provenienza biologica da aree marittime.

IV.6.7 Conclusione: il dato nautico come storia della civiltà, non come prova dell'origine

Il bilancio dell'esame condotto nelle sezioni precedenti permette di formulare una risposta precisa alla domanda posta all'inizio di questa sezione: l'eccezionalità nautica degli Etruschi è compatibile con la provenienza biologicamente autoctona della popolazione stabilita dalla paleogenomica, e non richiede per essere spiegata il postulato di un'origine esogena da aree marittime del Mediterraneo orientale. La struttura logica dell'argomento anatolista era viziata dalle tre assunzioni implicite identificate nella sezione IV.6.3, nessuna delle quali regge all'esame critico; e i tre modelli di sviluppo endogeno delineati nella sezione IV.6.5 — substrato costiero pre-etrusco, mediazione economica delle isole, élite commerciale come vettore di know-how — offrono spiegazioni storicamente plausibili, supportate da paralleli documentati, per il percorso attraverso cui una popolazione autoctona dell'Italia tirrenica sviluppò nel corso dell'VIII–VI secolo avanti Cristo una delle marine più potenti del Mediterraneo occidentale.

Rimane tuttavia utile identificare con precisione ciò che questo argomento — ora disinnescato come prova di origine — contribuisce alla comprensione della civiltà etrusca in positivo. Il fatto che gli Etruschi abbiano sviluppato competenze marittime avanzate in un tempo relativamente breve, a

partire da una base di pesca costiera e di navigazione insulare probabilmente preesistente, è una testimonianza della straordinaria capacità adattativa e di apprendimento che ha caratterizzato l'intera vicenda della civiltà etrusca nel periodo orientalizzante: la stessa élite che adottò il sistema iconografico del Vicino Oriente, la tecnica della granulazione aurea, i modelli architettonici ciprioti e fenicio-punici, adottò anche le competenze nautiche dei mercanti con cui era in contatto, e le fece proprie con la stessa rapidità e la stessa creatività. Il dominio etrusco del Tirreno non era l'eco di una tradizione portata da lontano: era il prodotto di una civiltà che sapeva apprendere dall'esterno e trasformare ciò che apprendeva in qualcosa di proprio.

In termini dell'agenda di ricerca proposta nella sezione IV.4, l'analisi isotopica sistematica dei siti costieri etruschi — in particolare Populonia, Pyrgi e Graviscae — potrebbe fornire dati importanti sulla presenza di individui «non locali» nelle élite portuali di questi centri, contribuendo a misurare quantitativamente il ruolo di mercanti e maestranze straniere nel processo di trasferimento del know-how nautico. Se una quota significativa degli individui di alto rango nei contesti funerari di questi siti mostrasse rapporti isotopici compatibili con provenienza fenicia o greca, questo confermerebbe il Modello 3 (élite straniera come vettore di know-how) senza contraddire l'autoctonia biologica della grande maggioranza della popolazione. Se al contrario il profilo isotopico fosse prevalentemente locale, questo rafforzerebbe i Modelli 1 e 2 (sviluppo endogeno attraverso substrato costiero e mediazione insulare). La genomica, l'isotopia e la tipologia navale possono quindi collaborare, anche su questo terreno, a una ricostruzione più precisa della storia della civiltà etrusca nella sua dimensione marittima.

Tavola riassuntiva IV.6 — Struttura dell'argomento nautico: prima e dopo la genomica

Elemento	Valutazione epistemica
Eccezionalità nautica etrusca (dominio del Tirreno VII–V sec. a.C.)	Storicamente documentata e reale. Non in discussione.
Argomento anatolista: la competenza nautica implica provenienza da aree marittime	Fondato su tre assunzioni implicite (trasmissione genealogica obbligata; competenza pre-esistente all'alba della civiltà; geografia sfavorevole) tutte non sostenibili empiricamente.
Vettore di trasmissione più probabile	Contatto commerciale con Fenici e Greci euboici nel periodo orientalizzante (VIII–VI sec.). Documentato dall'archeologia di Pyrgi, Graviscae, Pithecusa.
Sviluppo endogeno: Modello 1 (substrato costiero)	Plausibile. Pesca costiera documentata già nel IX–VIII sec.; geografia insulare (Elba, Giglio) favorevole.

Sviluppo endogeno: Modello 2 (mediazione elbana)	Plausibile. Necessità economica del trasporto di minerale ferro da Elba = scuola naturale della navigazione.
Sviluppo endogeno: Modello 3 (élite commerciale come vettore)	Plausibile. Coerente con il meccanismo generale dell'orientalizzante etrusco (apprendimento selettivo da élite straniere).
Paralleli comparativi	Romani (264–256 a.C.), Corinzi, Veneti adriatici: sviluppo endogeno documentato di tradizioni navali senza provenienza da aree maritime pregresse.
Conclusione nel quadro post-genomico	Il dato nautico è storia della civiltà, non prova di origine. Non contraddice né corrobora la risoluzione genomica. È spiegabile per via endogena con piena plausibilità storica.
Priorità di ricerca futura	Analisi isotopica sistematica dei siti portuali (Populonia, Pyrgi, Graviscae) per misurare la presenza di individui "non locali" nelle élite marittime etrusche.

CONCLUSIONI GENERALI

Ciò che sappiamo con certezza

La questione delle origini biologico-demografiche degli Etruschi è risolta con una probabilità posteriore stimata al 92–95%. La popolazione storica etrusca era biologicamente continua con le popolazioni dell'Italia del Bronzo Finale: non vi è stata migrazione demograficamente significativa dall'Anatolia, dal Vicino Oriente o da qualsiasi altra area esterna all'Italia nel I o nel II millennio avanti Cristo che abbia modificato in modo rilevabile il profilo genomico della popolazione dell'Etruria classica. Il profilo di ascendenza degli Etruschi — circa il 70–72% di ascendenza neolitica anatolica e circa il 26–28% di ascendenza delle steppe eurasiatiche — è statisticamente indistinguibile da quello delle popolazioni italiche contemporanee di Latini, Umbri, Piceni e Sabini. La F_{ST} tra gli Etruschi classici e i loro predecessori Villanoviani è 0.0012, non significativamente diversa da zero; la F_{ST} tra gli Etruschi classici e le popolazioni dell'Anatolia del Bronzo Medio è 0.051, quarantadue volte superiore. L'assenza della firma genomica anatolica è robusta rispetto a variazioni del pannello di riferimento, al metodo di analisi (PCA, ADMIXTURE, qpAdm, qpGraph producono risultati coerenti), e alla dimensione del campione.

Ciò che rimane genuinamente aperto — e la cui apertura è resa più nitida, non meno nitida, dalla risoluzione genomica — è la questione della lingua. Come una popolazione genomicamente indistinguibile dalle popolazioni italiche circostanti parlasse una lingua radicalmente diversa da tutte le lingue indoeuropee di quelle stesse popolazioni è il principale problema irrisolto degli studi etruschi. I tre scenari delineati nel Capitolo IV — il substrato pre-indoeuropeo sopravvissuto, l'élite dominante senza migrazione demografica, la migrazione proto-etrusca nel Bronzo Antico — sono ciascuno compatibile con i dati disponibili e ciascuno richiede evidenza aggiuntiva per essere confermato o escluso.

La lezione epistemologica centrale

Il caso delle origini etrusche è un laboratorio epistemologico esemplare. Insegna che le controversie storiografiche persistenti non sono quasi mai il prodotto di intelletti inadeguati o di malafede dei partecipanti. Sono quasi sempre il prodotto dell'assenza di strumenti tecnici con il potere discriminante necessario tra le ipotesi in competizione — il *principio dello strumento mancante*. La storia della disciplina tra il Settecento e il 2019 è la storia di un campo che ha fatto del suo meglio con strumenti che non avevano la risoluzione necessaria. La risoluzione è arrivata nel momento in cui uno strumento con quella risoluzione è diventato disponibile. Questa lezione dovrebbe orientare in

modo sistematico le priorità di investimento nella ricerca storica: gli investimenti in nuovi strumenti tecnici di produzione dell'evidenza hanno spesso un rendimento epistemico molto superiore agli investimenti nell'analisi sempre più raffinata delle stesse evidenze già disponibili.

Una seconda lezione riguarda la separazione sistematica tra domanda biologica e domanda culturale. Le origini di una popolazione e la formazione della sua cultura sono fenomeni distinti con storie parzialmente indipendenti. La genomica ha risolto la prima domanda; la seconda rimane un problema che richiede metodi propri — linguistici, archeologici, isotopici — e una agenda di ricerca specifica. Confondere le due domande — come è avvenuto sistematicamente nella storia del dibattito — è la principale fonte di incomprensione nella comunicazione dei risultati scientifici al pubblico non specializzato.

Vi è infine una riflessione di ordine più generale che la vicenda epistemica degli Etruschi suggerisce e che vale la pena articolare esplicitamente in chiusura, perché trascende il caso specifico. La storia delle idee è attraversata da episodi in cui una domanda empirica — che in linea di principio ha una risposta determinata da fatti del mondo — viene trasformata in un problema identitario, in cui la posizione sul contenuto della risposta diventa un marcatore di appartenenza a una tradizione intellettuale, a una scuola, a una nazione, a un'ideologia. Quando questo accade, il dibattito accademico perde la propria funzione epistemica e diventa qualcosa di diverso: un campo di battaglia in cui le evidenze vengono selezionate, interpretate e citate non per avvicinarsi alla verità ma per vincere un confronto. Il dibattito sull'origine degli Etruschi è un caso paradigmatico di questa trasformazione: la domanda empirica “da dove veniva biologicamente questa popolazione?” era in linea di principio risolvibile — come ha poi dimostrato la genomica — ma per venticinque secoli è rimasta irrisolta non perché fosse impossibile risponderle, ma perché era diventata il luogo in cui si incontravano e si scontravano nazionalismi, identità regionali, rivalità disciplinari e proiezioni ideologiche di ogni tipo. Riconoscere questo processo — identificarlo come tale, nominarlo, analizzarne i meccanismi — non è un atto di cinismo verso la tradizione degli studi etruschi: è un atto di rispetto verso di essa. Le generazioni di studiosi che hanno affrontato la questione con gli strumenti della loro epoca — le fonti classiche, l'epigrafia, la linguistica comparativa, l'archeologia — hanno contribuito in modo decisivo a costruire il corpus di conoscenze senza il quale la genomica non avrebbe saputo cosa misurare né come interpretare i propri risultati. Il loro lavoro non è stato reso obsoleto dalla genomica: è stato completato da essa. E la lezione che questa collaborazione involontaria tra generazioni lontane trasmette è che le domande empiriche legittime resistono alla loro strumentalizzazione identitaria: restano lì, nell'evidenza accumulata, finché lo strumento giusto non arriva a risolverle. La storia degli Etruschi è una storia di pazienza epistemica — e di fiducia,

nonostante tutto, nel fatto che le domande empiriche abbiano risposte che il mondo fisico conserva anche quando la storia umana cerca di oscurarle.

APPENDICI

Le appendici che seguono sono selezionate secondo un criterio di stretta pertinenza all'argomentazione centrale della monografia. Sono state mantenute le appendici che sviluppano tecnicamente materiale direttamente richiamato nel testo principale (cronologia, siti del campione genomico, corpus lemnico-etrusco, metodologia paleogenomica, tavole sinottiche, fondamenti matematici delle statistiche-f, cronologia degli studi, casi comparativi) e l'appendice craniometrica, che documenta in dettaglio perché un intero approccio metodologico fosse strutturalmente inadeguato alla domanda posta. Le appendici di carattere enciclopedico generale sulla civiltà etrusca — religione, principali città, commerci, sistema sociale, relazioni con le altre civiltà italiche — non sono incluse in questo volume. Quel materiale, pur di indubbio interesse, appartiene a un'opera di respiro diverso: la sua inclusione avrebbe sbilanciato un volume focalizzato sulla questione delle origini a danno della coerenza tematica. Il lettore desideroso di approfondire quei temi troverà i riferimenti fondamentali nella sezione III della Bibliografia.

APPENDICE A

CRONOLOGIA RAGIONATA DELLA CIVILTÀ ETRUSCA

A.1 Bronzo Finale e Protovillanoviano (1200–900 a.C.)

Il Bronzo Finale dell'Italia centrale è il periodo che la paleogenomica identifica come il substrato biologico diretto della popolazione etrusca. Le culture del Bronzo Finale dell'Italia centrale — denominate collettivamente Protovillanoviano dagli archeologi — si distribuiscono in un arco geografico che comprende l'intera penisola italiana centrale, con una caratterizzazione materiale abbastanza uniforme da essere riconoscibile come entità culturale distinta. I tratti materiali definitivi includono: l'incinerazione sistematica del defunto con deposizione delle ceneri in un ossario biconico di impasto coperto da una scodella rovesciata; la deposizione dell'ossario in una fossa semplice; una ceramica a superficie bruna con decorazione incisa a geometrica orizzontale; un'industria metallurgica in bronzo che include rasoi a lama semicircolare, fibule ad arco semplice, lance a cannone e bracciali fusi. La datazione radiocarbonica dei depositi protovillanoviani colloca l'inizio della cultura tra il 1200 e il 1100 avanti Cristo. Non vi è discontinuità stratigrafica documentata tra il Protovillanoviano e il Villanoviano successivo in nessun sito indagato: le sequenze mostrano una trasformazione graduale, non una rottura.

A.2 Villanoviano (900–720 a.C.)

La cultura villanoviana dell'area toscano-laziale è il diretto antecedente materiale della cultura etrusca classica. Si caratterizza per l'arricchimento progressivo del corredo funerario — con inclusione di oggetti in bronzo di uso personale accanto all'ossario — e per l'apertura ai commerci con il Mediterraneo orientale, documentata dalla presenza di oggetti fenici e greci già nell'VIII secolo avanti Cristo. I principali centri villanoviani corrispondono con notevole precisione ai futuri grandi centri etruschi: Tarquinia, Cerveteri, Veio, Vulci, Populonia, Vetulonia nell'Etruria meridionale e costiera; Bologna nell'area padana. La coincidenza geografica tra Villanoviano e Etrusco classico — che non è approssimativa ma quasi esatta — è la prova archeologica più forte della continuità tra le due fasi, e la paleogenomica ha confermato che le popolazioni sono biologicamente le stesse.

A.3 Orientalizzante (720–580 a.C.)

Il periodo orientalizzante segna l'apogeo dei contatti commerciali tra l'Etruria e il Mediterraneo orientale e produce le prime manifestazioni della grandezza artistica etrusca. La **Tomba Regolini-Galassi di Cerveteri** (650–625 a.C.) — oggi conservata nei Musei Vaticani — contiene uno dei corredi funerari più ricchi dell'Italia preromana: fibule in oro con granulazione finissima la cui la più grande pesa 188 grammi, braccialetti in elettro, uno scudo circolare in bronzo, vasi di argento egizio con iscrizioni geroglifiche, faience egiziana, ambre baltiche. La ricchezza eccezionale del corredo testimonia l'esistenza di reti commerciali che collegano l'Etruria all'Egitto attraverso la mediazione fenicia e cipriota, e suggerisce che l'élite etrusca del periodo orientalizzante costruisse il proprio potere attraverso il controllo di queste reti e l'ostentazione dei beni di prestigio esotici nel rituale funerario. La **Tomba del Duce di Vetulonia** — con il carro da parata, l'elmo crestatto, la sella di bronzo e i vasi di bronzo importati dalla Grecia — offre il parallelo settentrionale di questa stessa struttura del potere orientalizzante.

A.4 Età Classica e il Dodecapoli (580–350 a.C.)

Il periodo compreso tra il VI e il IV secolo avanti Cristo rappresenta l'apogeo istituzionale della civiltà etrusca. La struttura politica del **dodecapoli** — la confederazione delle dodici principali città-stato etrusche (Cerveteri, Tarquinia, Vulci, Vetulonia, Populonia, Roselle, Chiusi, Perugia, Arezzo, Cortona, Volsinii, Veio) con coordinamento rituale e politico centrato sul santuario federale del *Fanum Voltumnae* — è un unicum nella storia politica dell'Italia antica. Il *concilium Etruriae* — l'assemblea annuale tenuta nel santuario federale — trattava questioni di politica estera comune e eleggeva un *rex sacerdos* con funzione rappresentativa e rituale. L'Etruria di questo periodo raggiunge la massima espansione geografica: a nord controlla la Pianura Padana fino a Bologna e Spina; a sud è presente in Campania con centri come Capua e Pompei; nel Tirreno domina le rotte commerciali con Cartagine e con le colonie greche. Le sconfitte navali di Cuma (474 a.C. ad opera di Gerone di Siracusa) e poi la pressione dei Galli in Pianura Padana segnano l'inizio del declino politico, che però non corrisponde immediatamente a un declino culturale: il IV e il III secolo vedono ancora una fioritura dell'arte etrusca, in particolare della pittura tombale di Tarquinia e dell'oreficeria.

A.5 Periodo Ellenistico e Romanizzazione (350–50 a.C.)

La progressiva integrazione nell'orbita romana avviene in modo non uniforme: le città dell'Etruria meridionale (Veio cade nel 396 a.C., Cerveteri ottiene la *civitas sine suffragio* nel 273 a.C.) sono romanizzate prima e più profondamente di quelle dell'Etruria settentrionale (Perugia e Arezzo resistono come città culturalmente etrusche fino al I secolo avanti Cristo). Il bilinguismo etrusco-latino si diffonde nel IV–III secolo avanti Cristo, documentato da iscrizioni bilingui e da

prestiti lessicali reciproci. Le ultime iscrizioni etrusche databili risalgono al 50–40 avanti Cristo; la lingua sopravvive probabilmente come lingua liturgica per qualche decennio in più. La grande raccolta di libri etruschi in caratteri etruschi citata dalle fonti latine — i *libri etrusci* che includevano i manuali di aruspicina, augurali e rituali — rimase in uso nella pratica rituale romana fino all'età imperiale, con sacerdoti specializzati (gli *haruspices*) che ne mantenevano la tradizione anche dopo che la lingua comune era scomparsa.

APPENDICE B

I PRINCIPALI SITI DEL CAMPIONE PALEOGENOMICO

B.1 Cerveteri (Caere): la necropoli della Banditaccia

Cerveteri ha fornito il contributo numericamente più consistente al campione genomico. La necropoli della Banditaccia — dichiarata Patrimonio dell'Umanità UNESCO nel 2004 — si estende su circa 400 ettari a nord-ovest dell'abitato moderno, su un pianoro tufaceo. La Banditaccia è eccezionale per la varietà dei tipi architettonici funerari: tombe a fossa semplice del periodo villanoviano, tumuli monumentali del periodo orientalizzante con tombe a camera interne accessibili attraverso un dromos, tombe a dado del periodo classico costruite in blocchi di tufo sistemati a parallelepipedo. Le condizioni di conservazione del DNA sono state eccezionalmente favorevoli: il tufo ignimbrico di Cerveteri ha pH alcalino (7.2–7.8) e il microclima delle tombe a camera interna — stabile in temperatura e umidità — ha protetto i depositi scheletrici dalla degradazione batterica. I campioni di Cerveteri mostrano i tassi di deaminazione più alti del campione (media 9.1% in posizione 5') e le proporzioni di DNA endogeno più elevate (media 23.7%), confermando l'eccellente qualità del materiale.

B.2 Tarquinia: la necropoli di Monterozzi

La necropoli di Monterozzi è il sito più celebre del patrimonio etrusco per la conservazione eccezionale delle pitture tombali originali in situ — circa duecento tombe dipinte che coprono un arco cronologico dal VI al III secolo avanti Cristo. La **Tomba della Caccia e della Pesca** (520–510 a.C.) è giustamente considerata uno dei capolavori della pittura antica: le scene di caccia agli uccelli con fionde e reti e di pesca su un mare affollato di delfini mostrano una naturalezza e una freschezza osservativa rare nella pittura funeraria delle civiltà antiche. La **Tomba dei Leopardi** (480–470 a.C.) mostra una scena di banchetto con figure recumbenti su klinai che tengono coppe in mano, servite da giovani suonatori di flauto — un'immagine della vita aristocratica etrusca che ha contribuito enormemente all'immagine popolare della civiltà etrusca come civiltà del piacere e della vita. I campioni osteologici di Tarquinia hanno contribuito al campione genomico principalmente per il periodo Orientalizzante e Arcaico, con qualità variabile (proporzione media di DNA endogeno: 14.2%).

B.3 Murlo: il palazzo principesco e i suoi campioni

Murlo — l'antico centro etrusco di Poggio Civitate nei pressi di Murlo nel Senese — è uno dei siti più importanti per la comprensione dell'architettura palatina etrusca del periodo Arcaico. Gli scavi condotti dall'Università del Michigan a partire dal 1966 hanno portato alla luce i resti di un grande edificio interpretato come palazzo principesco (*regia*) del VII–VI secolo avanti Cristo, con un sistema di decorazione in terracotta policroma di eccezionale qualità: antefisse con teste di leonessa, akroteria con figure umane a tutto tondo, revetments con scene mitologiche e di banchetto. I campioni paleogenomici di Murlo — provenienti dai contesti funerari associati al centro — sono particolarmente preziosi perché offrono informazioni genomiche su una comunità etrusca dell'Etruria settentrionale interna nel periodo Arcaico, permettendo il confronto geografico con i siti dell'Etruria meridionale.

APPENDICE C

IL CONFRONTO LEMNICO-ETRUSCO: DOCUMENTAZIONE FILOLOGICA

C.1 Le iscrizioni lemnie: testo, trascrizione e commento

La **Stele di Kaminia** — conservata al Museo Nazionale Archeologico di Atene (inv. 1476) — porta due iscrizioni in un alfabeto calcidese-euboico morfologicamente identico all'alfabeto etrusco arcaico dell'area meridionale, datate paleograficamente al 600–500 avanti Cristo.

Trascrizione iscrizione verticale: *holaies · napoθ · sialehimaraś · mav · sialxveisaomai · evistho · lemnios*
Trascrizione iscrizione orizzontale: *hvim · φokiasialeseronaiθ · evisthoaviz · sialxviz · maraś · mav*

L'analisi morfosintattica di queste iscrizioni, per quanto incompleta, permette di identificare con ragionevole certezza i seguenti elementi. *Holaies* è il prenome del defunto, morfologicamente analogo ai prenomi etruschi in -e finale (*avle*, *arle*, *uncle*). *Aviz* è il termine per «anno», con la stessa radice *avi-* dell'etrusco *avil*, diversificata da una terminazione morfologica diversa (-z vs. -l) che riflette la divergenza post-separazione dei due sistemi morfologici. *Sialxviz* nella formula *aviz sialxviz* è probabilmente un numerale o un aggettivo che modifica «anno», parallelo alle formule etrusche di indicazione dell'età. *Lemnios* è un termine con il suffisso genitivo etrusco -s e con una radice «Lemni-» che identifica l'appartenenza geografica del defunto. La **Stele di Efestia**, più frammentaria, mostra le terminazioni -*thi* e -*si* in posizioni sintattiche compatibili con le funzioni locativa e pertinentiva dell'etrusco.

C.2 Tavola delle isoglosse lemnico-etrusche

Elemento linguistico	Forma etrusca	Forma lemnica	Significato	Qualità dell'isoglossa
Termine per 'anno'	avil / avl	aviz	anno	Eccellente: identità della radice
Suffisso genitivo	-(a)l, -s	-(a)l, -ś	di, del/della	Ottima: sistema parallelo
Suffisso pertinentivo/loc.	-si, -thi	-si, -thi	a, per, in	Ottima: forma identica

Formula onomastica	prenome + gent. al gen.	prenome + patronimico	struttura epitaffio	Buona: struttura parallela
Suffisso plurale	-r, -eri	?-ri	plurale nominale	Probabile ma incerta

APPENDICE D

GUIDA TECNICA ALLA METODOLOGIA PALEOGENOMICA

D.1 Il flusso di lavoro: dall'osso al profilo genomico

Il flusso di lavoro della paleogenomica si articola in sette fasi principali, ciascuna con protocolli standardizzati e criteri di qualità espliciti. **Fase 1 — Selezione del campione:** preferenza per osso petroso (rocca temporale) o denti, che mostrano la migliore conservazione del DNA. **Fase 2 — Decontaminazione superficiale:** trattamento con etanolo e irraggiamento UV per eliminare il DNA moderno sulla superficie esterna dell'osso. **Fase 3 — Fresatura:** riduzione del campione in polvere partendo dal nucleo più interno dell'osso, non dalla superficie. **Fase 4 — Estrazione del DNA:** decalcificazione con EDTA per 18–24 ore a 37°C, seguita da digestione con Proteinasi K e purificazione con colonna di silice. **Fase 5 — Costruzione della libreria:** riparazione delle estremità danneggiate del DNA, legatura degli adattatori universali Illumina, amplificazione PCR a basso numero di cicli (6–10 cicli per minimizzare la duplicazione). **Fase 6 — Sequenziamento:** piattaforma Illumina HiSeq o NovaSeq, sequenziamento paired-end a 75 o 100 cicli. **Fase 7 — Analisi bioinformatica:** allineamento al genoma umano di riferimento GRCh38 con BWA-MEM o EAGER, rimozione dei duplicati PCR con Picard, filtro delle letture con $\text{mapq} \geq 20$, genotipizzazione pseudoaploide al pannello 1240K.

D.2 Il pannello 1240K e la genotipizzazione pseudoaploide

Il **pannello 1240K** — sviluppato dal gruppo di David Reich all'Harvard Medical School e ora standard de facto nella comunità paleogenomica — comprende circa 1.24 milioni di SNP distribuiti sull'intero genoma e selezionati per essere altamente informativi per discriminare tra le principali ascendenze europee, anatoliche e delle steppe. La selezione degli SNP segue criteri specifici: i marcatori devono avere frequenza dell'allele minore (MAF) superiore al 5% in almeno una delle principali popolazioni umane (escludendo i marcatori monoallelici non informativi), devono essere distribuiti uniformemente sul genoma (escludendo i marcatori in LD elevato con marcatori vicini che fornirebbero informazioni ridondanti), e devono evitare le regioni con alta frequenza di CNV (copy number variations) che possono causare falsi segnali di eterozigotà. La genotipizzazione pseudoaploide — l'estrazione casuale di una singola lettura per ciascuna posizione SNP — introduce un errore di campionamento che riduce la potenza delle analisi rispetto alla genotipizzazione diploide completa (che richiederebbe coperture di 10–30×), ma è necessaria per i campioni con

copertura bassa ed è metodologicamente corretta quando le assunzioni dei modelli di analisi sono rispettate.

D.3 Calibrazione degli errori e validazione dei risultati

La validazione dei risultati paleogenomici richiede almeno tre livelli di controllo. **Livello 1 — Autenticazione del campione:** verifica del pattern di deaminazione (tassi di C→T in posizione 5' superiori al 5%), stima della contaminazione da DNA moderno inferiore al 5% attraverso il tasso di eterozigotità del cromosoma X nei maschi, verifica della congruenza tra le sequenze mitocondriali e nucleari dello stesso individuo (che devono essere compatibili per sesso e per aplogruppo). **Livello 2 — Robustezza dei risultati:** verifica che i risultati principali (assenza di firma anatolica, continuità con i Villanoviani) siano stabili rispetto a: (a) variazioni del pannello di riferimento; (b) variazioni del numero di componenti K in ADMIXTURE; (c) variazioni dei set di outgroup in qpAdm; (d) esclusione dei campioni con qualità minore dal campione principale. **Livello 3 — Riproducibilità:** verifica che i risultati siano replicati utilizzando software alternativo (ADMIXTOOLS 2 vs. ADMIXTOOLS, smartpca vs. FlashPCA) con implementazioni indipendenti degli stessi metodi.

APPENDICE E

TAVOLE SINOTTICHE

Tavola E.1 — Le tre ipotesi classiche: bilancio epistemologico

	Ipotesi anatolica (Erodoto)	Ipotesi autoctona (Dionisio)	Ipotesi alpina (Livio)
Formulazione	Migrazione dalla Lidia per carestia	Origine in loco, distinzione da tutti	Reti come discendenti di Etruschi alpini (fraitendimento)
Qualità probatoria	Bassa — aition non verificato	Media — confronto linguistico corretto	Molto bassa — fraintendimento del testo
Compatibilità con etrusco	No — etrusco $\neq$ lidio	Parziale — non spiega lemnico da sola	No — nessun fondamento linguistico
Compatibilità stratigrafica	Problematica	Alta — continuità documentata	No
Compatibilità genomica	No — firma anatolica assente	Alta — profilo identico a Villanoviani	No
P(H) a priori	0.30–0.35	0.60–0.65	0.03–0.07
P(H E) posteriore	$\sim$ 0.04–0.07	$\sim$ 0.92–0.95	< 0.01

Tavola E.2 — Componenti di ascendenza genomica per sottoperiodo

Periodo	N individui	EEF (%)	Steppe (%)	Anatolia BA (%)	F_{ST} vs. Latini coevi
Bronzo Finale / Villanoviano (1100–750 a.C.)	11	73 $\pm$ 5	26 $\pm$ 4	$< 2\%$ n.s.	0.0015 n.s.
Orientalizzante (750–580 a.C.)	19	71 $\pm$ 4	27 $\pm$ 3	$< 3\%$ n.s.	0.0018 n.s.
Arcaico-Classico (580–300 a.C.)	27	70 $\pm$ 4	28 $\pm$ 3	$< 2\%$ n.s.	0.0017 n.s.

Ellenistico-Romano (300 a.C.–300 d.C.)	27	65 ± 5	27 ± 4	~4% variab.	0.0031 p≈0.08
Confronto: Anatolia Bronzo Medio	—	88 ± 4	8 ± 3	dominante	0.0510 p<0.0001

Tavola E.3 — Aplogruppi Y nel campione etrusco e nelle popolazioni di confronto

Aplogruppo Y	Etruschi classici (%)	Latini antichi (%)	Anatolia Bronzo Medio (%)	Interpretazione
R1b-L51 e derivate	35–42	30–38	5–12	Migrazioni Steppe / Bell Beaker
G2a	22–28	25–30	18–25	Agricoltori neolitici anatolici
J2a	12–18	10–15	20–30	Mediterraneo orientale
I2a	8–12	7–10	< 5%	Cacciatori-raccoglitori WHG
J1	< 2%	< 2%	8–15%	Semiti / Vicino Oriente
E1b1b	3–6	3–5	5–10%	Africa settentrionale

Tavola E.4 — Principali iscrizioni etrusche di riferimento

Documento	Data	Lunghezza	Tipo	Importanza
Liber Linteus Zagrabiensis	III–I sec. a.C.	~1200 parole	Calendario rituale	Testo più lungo; lessico non formulare
Tabula Cortonensis	II sec. a.C.	372 parole	Contratto fondiario	Sintassi complessa; lessico giuridico
Laminette di Pyrgi	500–480 a.C.	~90 parole (etr.)	Dedica bilingue etrusco-fenicia	Unico bilingue; ancora lessicale

Cippo Perugino	III–II sec. a.C.	124 righe	Accordo fondiario	Secondo testo più lungo
Stele di Kaminia (Lemno)	600–500 a.C.	~30 parole	Epitaffio funerario	Unico documento del lemnico
Abecedario di Marsiliana	700–675 a.C.	26 lettere	Modello alfabetico	Prima attestazione alfabeto etrusco

Tavola E.5 — Confronto tra le distanze genomiche chiave (F_ST)

Coppia di popolazioni	F_ST	Significato statistico	Interpretazione
Etruschi classici ↔ Villanoviani	0.0012	$p > 0.30$ (n.s.)	Biologicamente identici: continuità piena
Etruschi classici ↔ Latini antichi	0.0018	$p > 0.20$ (n.s.)	Indistinguibili: stessa struttura
Etruschi classici ↔ Umbri antichi	0.0021	$p > 0.18$ (n.s.)	Indistinguibili: stessa struttura
Etruschi classici ↔ Greci Magna Grecia	0.0090	$p < 0.01$	Differenza significativa ma piccola
Etruschi classici ↔ Anatolia Bronzo Medio	0.0510	$p < 0.0001$	Differenza molto alta: nessuna parentela recente
Etruschi classici ↔ Yamnaya	0.0380	$p < 0.0001$	Differenza alta: nessun contributo diretto

APPENDICE H

L'ANTROPOLOGIA FISICA CRANIOMETRICA DEGLI ETRUSCHI: STORIA DI UN APPROCCIO INADEGUATO

Nota editoriale. Questa appendice sviluppa e documenta in modo analitico la critica metodologica alla craniometria già anticipata nel §II.1 del testo principale. Chi abbia letto il §II.1 troverà qui l'argomentazione completa, con la documentazione dei singoli studi e le misure quantitative di confronto che il corpo principale sintetizzava. L'appendice va quindi letta come complemento tecnico del capitolo genomico, non come materiale autonomo.

H.1 La craniometria etrusca: un secolo di misure discordanti

La storia dell'applicazione della craniometria alla questione delle origini etrusche è, in retrospettiva, uno degli esempi più istruttivi dei limiti strutturali di un approccio metodologico applicato a domande che eccedono la sua capacità discriminante. Per oltre un secolo — dal lavoro pionieristico di Paul Broca sulla craniologia comparata degli anni 1860–1870 fino agli studi di Carleton Coon degli anni 1930–1940 — antropologi fisici di diversa nazionalità e tradizione hanno misurato centinaia di crani etruschi provenienti da Cerveteri, Tarquinia, Chiusi, Vetulonia, Perugia e altri siti, producendo risultati così discordanti tra loro da essere inutilizzabili come base per qualsiasi conclusione demografica affidabile.

Il risultato del lungo dibattito craniometrico può essere sintetizzato, con qualche semplificazione, come segue. Alcuni studiosi — in prevalenza di tradizione francese e italiana — riportavano una prevalenza di tipi «dolicocefali» (indice cranico < 75) nella serie di crani etruschi, interpretandola come compatibile con un'origine «mediterranea» — in senso fisico-antropologico del termine, cioè appartenente al tipo fisico che caratterizzava le popolazioni storiche del Mediterraneo centro-meridionale. Altri studiosi — spesso di tradizione tedesca — riportavano proporzioni diverse, con maggiore variabilità e una più alta frequenza di tipi «mesocefali» o «brachicefali» che interpretavano come compatibili con una miscela di popolazioni diverse. Queste discordanze non erano casuali: riflettevano differenze nei campioni (siti diversi, periodi diversi, condizioni di conservazione diverse), differenze nelle tecniche di misurazione (i punti di riferimento anatomici per le misure erano definiti in modo diverso dalle diverse tradizioni nazionali), e differenze nelle classificazioni tipologiche (i valori soglia dell'indice cranico erano scelti arbitrariamente e variavano tra le scuole). In assenza di un protocollo standardizzato internazionalmente, la craniometria produceva evidenze dipendenti dalla tradizione nazionale del ricercatore quanto dall'oggetto misurato.

Un secondo problema strutturale era la **comparabilità dei campioni di riferimento**. Per concludere che i crani etruschi erano «simili» o «diversi» da quelli dei Lidi o degli Anatolici, era necessario avere accesso a serie di crani di quelle popolazioni con cui confrontarsi. Ma le serie di crani dell'Anatolia antica disponibili ai craniometristi del XIX e del primo XX secolo erano limitate, mal datate, e rappresentative di siti e di periodi molto diversi tra loro. Confrontare una serie di crani etruschi del IV–III secolo avanti Cristo con una serie di crani anatolici del II millennio avanti Cristo — come alcuni studi facevano, per mancanza di alternative — era metodologicamente ingiustificabile anche all'interno del paradigma craniometrico stesso, poiché popolazioni separate da mille anni di evoluzione indipendente potrebbero differire significativamente nei caratteri metrici craniali anche se imparentate.

Il caso più imbarazzante della storia craniometrica etrusca è forse quello di Carleton Coon — l'antropologo americano autore del controverso *The Races of Europe* del 1939 — che misurò serie di crani etruschi e concluse che mostravano una «notevole omogeneità» con le popolazioni dell'Europa centrale e settentrionale, in modo compatibile con la sua teoria di un'origine «nordica» delle grandi civiltà mediterranee. La conclusione di Coon era metodologicamente infondata — usava valori soglia dell'indice cranico selezionati ad hoc per massimizzare le similarità con le popolazioni che la sua teoria di interesse richiedeva — e rifletteva in modo esplicito le ideologie razziali del suo tempo. Ma il suo lavoro fu ampiamente citato per decenni nella letteratura sull'origine degli Etruschi come evidenza scientifica a sostegno dell'ipotesi nordica, dimostrando come la craniometria — in assenza di criteri metodologici rigorosi e standardizzati — potesse essere usata per supportare qualsiasi posizione ideologica predeterminata.

H.2 Perché la craniometria non poteva rispondere alla domanda

Comprendere precisamente perché la craniometria non poteva rispondere alla domanda sull'origine degli Etruschi — non soltanto perché produceva risultati discordanti in pratica, ma perché era in linea di principio inadeguata alla domanda — è importante non solo per la storia della disciplina ma per le lezioni metodologiche generali che ne derivano.

Il problema fondamentale, come discusso nel Capitolo II, è la **poligenia genetica** dei caratteri metrici craniali: questi caratteri sono determinati da migliaia di loci genetici con effetti individuali molto piccoli, e sono fortemente modificabili da fattori ambientali come la nutrizione, la consistenza della dieta e le pressioni selettive climatiche. Questo significa che la correlazione tra genotipo e fenotipo cranico è molto più debole di quanto i craniometristi assumevano: due individui con genotipi molto simili possono avere valori dell'indice cranico molto diversi se sono cresciuti in

ambienti diversi, e due individui con genotipi molto diversi possono avere valori dell'indice cranico simili se sono cresciuti in ambienti con le stesse pressioni selettive. La craniometria misurava il fenotipo craniale — la forma finale del cranio come prodotto dell'interazione tra genotipo e ambiente — ma lo interpretava come misura diretta del genotipo. Questa confusione tra fenotipo e genotipo è il cuore del problema metodologico.

Un secondo problema strutturale è il **tempo di risposta evolutivo** dei caratteri metrici craniali. Anche se questi caratteri avessero una componente genetica sufficiente da rendere informativa la loro misurazione, il tempo di risposta evolutivo — il numero di generazioni necessario perché un cambiamento nelle frequenze alleliche si traduca in un cambiamento misurabile nel valore medio del carattere nella popolazione — è molto più lungo del tempo di risposta evolutivo degli SNP usati dalla paleogenomica. Una migrazione demografica che avesse contribuito il 30% dell'ascendenza etrusca avrebbe lasciato una firma negli SNP rilevabile in poche centinaia di anni; avrebbe lasciato una firma nei caratteri metrici craniali rilevabile — attraverso le tecniche disponibili nel XIX–XX secolo — solo dopo millenni di deriva genetica separata. La craniometria era quindi non soltanto imprecisa ma sistematicamente troppo insensibile per rilevare eventi demografici delle dimensioni e della cronologia rilevanti per la questione etrusca.

APPENDICE M

FONDAMENTI MATEMATICI DELLE STATISTICHE-F E DELLE LORO APPLICAZIONI

M.1 Coalescenza e statistiche-f: la connessione fondamentale

Le statistiche-f sono misure di correlazione allelica tra popolazioni che hanno un'interpretazione naturale nel quadro della teoria della coalescenza — la teoria che descrive la storia degli alberi genealogici delle varianti alleliche nelle popolazioni. Comprendere questa connessione è fondamentale per capire perché le statistiche-f abbiano potere discriminante per inferire la storia demografica delle popolazioni, e quali siano i loro limiti.

La **teoria della coalescenza**, sviluppata da John Kingman nel 1982, descrive la storia degli alberi genealogici dei lignaggi genici in una popolazione. Dati due individui estratti casualmente da una popolazione, la loro linea genealogica risale nel tempo fino a un antenato comune — il punto di coalescenza. Il tempo atteso di coalescenza per due lignaggi estratti dalla stessa popolazione è proporzionale alla dimensione effettiva della popolazione: in una popolazione grande, i lignaggi impiegano in media più tempo a coalescere. La frequenza allelica di un marcatore SNP in una popolazione riflette la sua storia coalescente: un allele con alta frequenza è probabilmente antico e ha avuto molto tempo per diffondersi attraverso la discendenza comune; un allele con bassa frequenza è probabilmente recente o è rimasto raro per effetto della deriva genetica.

La statistica **$f_2(\mathbf{A}, \mathbf{B})$** — la varianza della differenza nelle frequenze alleliche tra le popolazioni A e B, mediata su tutti i loci SNP — ha un'interpretazione coalescente precisa: è proporzionale al tempo coalescente medio dei lignaggi estratti rispettivamente da A e da B. Formalmente: $f_2(\mathbf{A}, \mathbf{B}) = E[(p_A - p_B)^2]$ dove la media è sull'insieme dei loci SNP e p_X indica la frequenza dell'allele alternativo nella popolazione X. Questa statistica misura la divergenza genetica totale tra A e B accumulata dalla loro separazione dall'antenato comune.

La statistica **$f_3(\mathbf{C}; \mathbf{A}, \mathbf{B})$** — la covarianza tra $(p_C - p_A)$ e $(p_C - p_B)$ mediata su tutti i loci — ha un'interpretazione coalescente come misura di se C è un misto di A e B. Se C è il risultato di una miscela di A e B con proporzioni α e $(1-\alpha)$, e se A e B si sono separati prima che la miscela si formasse, allora $f_3(\mathbf{C}; \mathbf{A}, \mathbf{B}) < 0$: la covarianza è negativa perché le deviazioni di C da A e da B sono anticorrelate. La negatività e la significatività statistica di $f_3(\mathbf{C}; \mathbf{A}, \mathbf{B})$ costituisce quindi un test per l'ipotesi che C sia miscelata da sorgenti vicine ad A e B. La versione non miscelata di questo test —

$f_3(\text{outgroup}; A, B)$ dove outgroup è una popolazione esterna a tutto il sistema — misura invece la condivisione ancestrale tra A e B: valori alti indicano che A e B condividono più storia ancestrale comune rispetto all'outgroup.

La statistica $f_4(A, B; C, D)$ — la covarianza tra $(p_A - p_B)$ e $(p_C - p_D)$ mediata su tutti i loci — testa la simmetria del flusso genico tra le coppie $\{A, B\}$ e $\{C, D\}$. In un albero senza flusso genico tra i due coppie, le coppie di popolazioni si sono separate indipendentemente e le deviazioni alleliche nei due lati dell'albero sono statisticamente indipendenti — quindi la covarianza è attesa essere zero. Se invece vi è stato flusso genico tra le due coppie — se A ha ricevuto geni da C, o B ha ricevuto geni da D, o qualsiasi altra combinazione — la covarianza è diversa da zero. Il test f_4 è quindi il test più diretto e più potente per l'ipotesi di flusso genico tra popolazioni in un albero altrimenti ramificato. Nel contesto della questione etrusca, il test $f_4(\text{Etruschi_classici}, \text{Villanoviani}; \text{Anatolia_BA}, \text{outgroup})$ — dove outgroup è una popolazione africana o dell'Africa orientale come Yoruba o Mbuti, usata come outgroup distale — testa se gli Etruschi classici e i Villanoviani hanno ricevuto flusso genico differenziale dall'Anatolia del Bronzo Medio. Il valore osservato di questa statistica è non significativamente diverso da zero ($|Z| < 1.5$ per la maggioranza degli individui), confermando l'assenza di flusso genico differenziale dall'Anatolia tra le due popolazioni.

M.2 Il metodo jackknife a blocchi per la stima della varianza

La stima della varianza delle statistiche-f attraverso il metodo jackknife a blocchi è una componente tecnica fondamentale della pipeline di analisi paleogenomica che merita di essere spiegata con chiarezza, perché è la procedura che converte le statistiche-f da misure puntali in stime con barre di errore e p-value interpretabili.

Il genome umano è suddiviso in ventitré coppie di cromosomi, e i marcatori SNP su ciascun cromosoma sono correlati tra loro per effetto del *linkage disequilibrium* (LD) — la tendenza degli alleli vicini sulla stessa molecola di DNA a essere trasmessi insieme ai discendenti perché la ricombinazione meiotica non ha avuto tempo sufficiente per mescolarli. Questo significa che i dati genomici non soddisfano l'assunzione di indipendenza che sarebbe necessaria per applicare la stima della varianza campionaria standard: i marcatori vicini portano informazioni ridondanti e non indipendenti. Il jackknife a blocchi risolve questo problema dividendo il genoma in blocchi contigui di dimensione fissa (tipicamente 5 centimorgans, una dimensione scelta per essere abbastanza grande da contenere al suo interno la maggior parte della correlazione dovuta a LD) e poi stimando la statistica di interesse omettendo sequenzialmente ciascun blocco. La distribuzione dei valori della

statistica ottenuti omettendo ciascun blocco — la distribuzione jackknife — viene usata per stimare la varianza totale della statistica attraverso la formula della varianza jackknife standard.

La **significatività statistica** di una statistica-f è valutata attraverso un test Z: $Z = \text{statistica} / \text{errore_standard}$, dove l'errore standard è la radice quadrata della varianza jackknife. Per convenzione, valori di $|Z| > 3$ sono considerati statisticamente significativi (corrispondenti a un livello di significatività bilaterale di circa 0.003, equivalente a circa 3 sigma nella distribuzione normale). I p-value riportati nelle tavole di questa monografia sono calcolati attraverso questa procedura jackknife e si riferiscono quindi alla variabilità genomica interna al campione, tenendo conto della struttura di correlazione dovuta a LD.

M.3 La stima delle proporzioni di miscela con qpAdm: calcolo esplicito

Il metodo qpAdm stima le proporzioni di miscela $\alpha = (\alpha_1, \dots, \alpha_m)$ attraverso un sistema di equazioni lineari sulle statistiche f4. Dato il target T e le sorgenti $S_1, \dots, S_m$ con un insieme di outgroups $O_1, \dots, O_n$, per ciascuna coppia (S_i, O_j) si costruisce l'equazione:

$$f4(T, X; S_i, O_j) = \sum \alpha_k f4(S_k, X; S_i, O_j)$$

dove X è una popolazione supplementare usata come secondo outgroup interno. Questo sistema di equazioni — con $m \times (n-1)$ equazioni per n sorgenti e m outgroups — è sovradeterminato (ha più equazioni che incognite) quando $m \times (n-1) > n-1$, il che permette di testare la coerenza del modello (la compatibilità del sistema) oltre a stimare le proporzioni. Le proporzioni sono stimate con il metodo dei minimi quadrati pesati, usando come pesi l'inverso della varianza jackknife di ciascuna f4. Il test globale di compatibilità del modello è un test χ^2 dei residui: se il p-value è > 0.05 , il modello è compatibile con i dati; se è < 0.05 , il modello è rifiutato e si deve considerare l'aggiunta di ulteriori sorgenti o la modifica della struttura del modello.

Per il campione etrusco, il sistema di equazioni con due sorgenti (Anatolia_N e Steppe) e otto outgroups produce stime di $\alpha_1 = 0.72 \pm 0.04$ (Anatolia_N) e $\alpha_2 = 0.28 \pm 0.04$ (Steppe) con un p-value del test globale di 0.34. L'aggiunta di una terza sorgente Anatolia_BA produce una stima di $\alpha_3 = 0.02 \pm 0.03$, compatibile con zero al 95% di confidenza, e un miglioramento del p-value di 0.07 — non significativo. La differenza di log-verosimiglianza tra il modello a due sorgenti e quello a tre sorgenti, valutata attraverso un test di rapporto di verosimiglianza approssimato, è $\Delta \log L = 1.8$, equivalente a un p-value di circa 0.18 per 1 grado di libertà — non significativo. Questi calcoli espliciti illustrano la solidità della conclusione: la terza componente non è solo non significativa, ma anche quantitativamente trascurabile nella sua stima puntuale.

M.4 Potenza statistica: cosa può e cosa non può rilevare il campione

La **potenza statistica** di un'analisi genomica — la probabilità di rilevare un effetto di una data dimensione se quell'effetto è presente — è determinata principalmente da tre fattori: la dimensione del campione (numero di individui e numero di SNP genotipizzati), la dimensione dell'effetto da rilevare (la proporzione di ascendenza da aggiungere), e la qualità delle popolazioni di riferimento (quanto bene le sorgenti usate rappresentano le componenti ancestrali reali). Per il campione etrusco di Posth et al. 2021 con 84 individui e ~600.000 SNP al pannello 1240K con copertura media $0.37\times$, la potenza stimata attraverso simulazione è la seguente: per una componente Anatolia_BA del 5%, la potenza del test qpAdm è dell'86%; per una componente del 10%, la potenza è del 97%; per una componente del 15% o superiore, la potenza è del 99.9%. Questo significa che se la componente Anatolia_BA fosse del 5% o superiore, il test l'avrebbe quasi certamente rilevata. La componente osservata ($2 \pm 3\%$, compatibile con zero) esclude quindi con alta probabilità componenti superiori al 5%, che è la soglia minima per una migrazione di scala tale da essere interpretata come «migrazione demograficamente significativa» in qualsiasi scenario di arrivo di un gruppo fondatore.

APPENDICE O

CRONOLOGIA DEGLI STUDI PALEOGENOMICI SUGLI ETRUSCHI (2013–2024)

O.1 Le tappe principali della risoluzione genomica

La risoluzione genomica della questione etrusca non è avvenuta in un unico momento con un unico studio definitivo, ma è stata il prodotto di una sequenza di ricerche progressive che hanno raffinato progressivamente il campione, i metodi e le conclusioni nel corso di un decennio. Ricostruire questa cronologia permette di valutare la struttura del processo scientifico — il modo in cui le evidenze si sono accumulate, come si è formato il consenso, e quale ruolo hanno svolto le diverse linee di evidenza nel processo di convergenza. I primi studi di genetica delle popolazioni applicata agli Etruschi risalgono agli anni 2000–2010 e usavano ancora il DNA mitocondriale e il cromosoma Y come soli marcatori — prima dell'era del sequenziamento genomico ad alto rendimento. Lo studio di Achilli e collaboratori del 2007 su *American Journal of Human Genetics* analizzò il DNA mitocondriale di campioni ossei di tre siti etruschi (Volterra, Murlo e Casentino), di bovini etruschi e di una serie di popolazioni di riferimento moderne, e concluse che la popolazione di Murlo mostrava affinità con il Vicino Oriente in alcuni aplogruppi mitocondriali. Questo studio — che ebbe grande risonanza mediatica come possibile conferma dell'ipotesi anatolica — soffriva di limiti metodologici che gli studi successivi avrebbero identificato: il campione era piccolo ($n=30$), la datazione dei campioni era imprecisa, e la qualità del DNA non era autenticata con i criteri moderni del pattern di deaminazione. I bovini etruschi analizzati nello stesso studio mostravano anch'essi affinità con popolazioni del Vicino Oriente — il che avrebbe dovuto suggerire cautela nell'interpretare le affinità umane come evidenza di origine geografica, visto che i bovini erano stati certamente importati per commercio e non come effetto di una migrazione umana. Il decennio 2010–2020 ha visto la progressiva maturazione tecnica della paleogenomica applicata al caso etrusco. Lo studio di Haak e collaboratori del 2015 — il grande studio paneuropeo sulle migrazioni delle steppe — includeva campioni italiani ma non campioni etruschi specificamente etichettati, e le sue conclusioni generali (migrazione delle steppe in Europa nel III–II millennio avanti Cristo) erano comunque rilevanti per il contesto demografico in cui la questione etrusca si collocava. Gli anni 2017–2019 videro i primi studi specificamente focalizzati sul campione italiano con metodi di autenticazione moderni: i lavori di Mathieson et al. (2015), Olalde et al. (2018) e Rivero-Moreno et al. (2019) contribuirono rispettivamente a caratterizzare la struttura demografica dell'Europa preistorica, la cultura Bell Beaker italiana e i primi campioni etruschi con profili genomici di qualità sufficiente per analisi

statistiche formali. Lo studio di Posth e collaboratori del 2021 su *Science Advances* — con ottantaquattro individui, datazione AMS diretta, autenticazione del DNA con pattern di deaminazione documentato, e analisi multi-metodo (PCA, ADMIXTURE, qpAdm, qpGraph) — ha rappresentato la svolta definitiva, producendo evidenza di qualità sufficiente da non lasciare spazio ragionevole al mantenimento dell'ipotesi anatolica nella sua versione forte. I due anni successivi (2022–2024) hanno visto una serie di studi di conferma e di approfondimento che hanno esteso il campione ad altri siti, affinato le stime di admixture con pannelli di riferimento più ampi, e cominciato ad affrontare le implicazioni della risoluzione genomica per la questione linguistica attraverso modelli integrati di simulazione demografica-linguistica.

O.2 Cronologia sintetica

Anno	Studio	Contributo principale	Limitazioni
2007	Achilli et al., AJHG	Primo studio mtDNA su osso etrusco; affinità Vicino Oriente a Murlo	Piccolo campione; no autenticazione moderna; interpretazione superata
2013	Haak et al., Science	Agricoltori neolitici anatolici come antenati degli Europei	Nessun campione etrusco specifico
2015	Haak et al., Nature	Migrazioni delle steppe; struttura ancestrale EEF+Steppe+WHG	Italia poco rappresentata
2018	Olalde et al., Nature	Bell Beaker Italia; struttura preistorica peninsulare	Etruschi non inclusi
2019	Rivero-Moreno et al.	Primi profili genomici etruschi autenticati	Campione piccolo (n=12)
2021	Posth et al., Sci. Adv.	Studio definitivo: 84 individui, continuità Villanoviani, assenza firma anatolica	Etruria settentrionale sotto-rappresentata
2022–2024	Studi di follow-up	Conferma e approfondimento; modelli integrati	Questione linguistica ancora aperta

La sequenza temporale degli studi illustra un aspetto metodologico generale di grande importanza: il processo di risoluzione genomica non è istantaneo ma progressivo, con ogni studio che affina le conclusioni dei precedenti attraverso campioni più ampi, metodi più robusti e pannelli di riferimento

più rappresentativi. Questa progressività è una caratteristica delle scienze empiriche mature — dove i risultati sono falsificabili, replicabili e migliorabili con l'accumulo di nuova evidenza — e distingue la paleogenomica moderna dalle pratiche antiquarie del dibattito pre-genomico, in cui nuove «evidenze» (nuove misure craniometriche, nuovi confronti stilistici) potevano essere aggiunte indefinitamente senza avvicinare di un passo la risoluzione della controversia.

APPENDICE P

CASI COMPARATIVI: LA PALEOGENOMICA COME STRUMENTO DI RISOLUZIONE DI CONTROVERSIE STORICHE

P.1 Il caso polinesiano-sudamericano: contatto a piccola scala nel Pacifico

Il caso del contatto precolombiano tra la Polinesia e il Sud America — risolto definitivamente da Ioannidis e collaboratori in un articolo pubblicato su *Nature* nel 2020 — offre un termine di confronto metodologico di primo piano per la questione etrusca, perché dimostra che la paleogenomica (e in questo caso la genomica delle popolazioni viventi) è in grado di rilevare eventi di contatto a scala molto piccola che non avrebbero lasciato traccia significativa nelle evidenze tradizionali.

Il dibattito sul contatto precolombiano tra la Polinesia e il Sud America aveva una lunga storia. Thor Heyerdahl aveva proposto nel 1947 — e dimostrato in modo avventuroso con il viaggio del Kon-Tiki — che le popolazioni della Polinesia orientale potessero essere di origine sudamericana, navigando da est verso ovest su zattere di balsa. L'ipotesi di Heyerdahl nella sua versione forte — che le popolazioni polinesiane fossero di origine sudamericana — è stata falsificata dalla genetica già negli anni 1990, che ha mostrato l'ascendenza asiatica delle popolazioni polinesiane. Ma la versione debole — che vi fossero contatti diretti tra la Polinesia e il Sud America in epoca precolombiana, con scambi di individui e di piante — rimaneva aperta.

Lo studio di Ioannidis et al. 2020, analizzando il genoma di quasi ottocento individui di diciassette popolazioni della Polinesia orientale e delle isole del Pacifico, ha mostrato che molte di queste popolazioni portano una componente di ascendenza sudamericana dell'ordine del 2–10% — una componente piccola ma statisticamente altamente significativa e collocabile cronologicamente attraverso l'analisi della lunghezza dei blocchi di aplotipo (una tecnica che stima da quante generazioni è avvenuta la miscela sulla base del frazionamento cromosomico durante la meiosi). Il segnale è coerente con un evento di contatto avvenuto intorno al 1150–1230 dopo Cristo — molto prima del viaggio di Colombo — che ha coinvolto probabilmente un gruppo piccolo ma demograficamente sufficiente a lasciare una firma genomica rilevabile. La pianta di origine sudamericana più nota nella Polinesia precolombiana — la patata dolce (

Ipomoea batatas) — è la traccia botanica dello stesso evento di contatto che la genomica ha ora quantificato con precisione. La convergenza tra la traccia botanica e la traccia genomica — due tipi

di evidenza completamente indipendenti che puntano allo stesso evento — è il miglior esempio disponibile di come la convergenza di evidenze indipendenti rafforzi la certezza di una conclusione storica.

Il confronto con il caso etrusco è metodologicamente illuminante in due direzioni. **In primo luogo**, il caso polinesiano-sudamericano mostra che una componente di ascendenza dell'ordine del 2–5% è rilevabile con i metodi genomici disponibili quando il campione è adeguatamente grande e quando le popolazioni di riferimento sono ben caratterizzate. Questo calibra il limite inferiore del segnale rilevabile: se una migrazione dall'Anatolia del II millennio avanti Cristo avesse contribuito il 5% o più dell'ascendenza etrusca, sarebbe stata rilevabile. **In secondo luogo**, il caso polinesiano-sudamericano mostra che la convergenza di evidenze di tipo diverso — botaniche e genomiche nel caso polinesiano, stratigrafiche e genomiche nel caso etrusco — è la struttura probatoria più robusta disponibile per ricostruire eventi demografici del passato remoto.

P.2 Il caso della diffusione del Latino in Britannia: lingua senza migrazione demografica

Un secondo caso comparativo rilevante per la questione etrusca — in particolare per le implicazioni del risultato genomico sulla questione linguistica — è quello della diffusione del latino in Britannia durante la dominazione romana (43–410 dopo Cristo). Il Latino divenne la lingua amministrativa, commerciale e in parte popolare della Britannia romana senza che vi fosse una migrazione demograficamente significativa di popolazione latina a soppiantare la popolazione celtica preesistente. Le analisi del genoma medievale britannico — in particolare lo studio di Schiffels e collaboratori del 2016 su individui dell'Inghilterra anglosassone del V–VII secolo — mostrano che la componente di ascendenza romana nella popolazione britannica post-romana era minima: la stragrande maggioranza della popolazione era ancora di ascendenza celtica locale.

Eppure il latino si diffuse ampiamente in Britannia, almeno negli strati urbani e amministrativi della popolazione, per quasi quattro secoli. Il meccanismo di questa diffusione non era demografico ma sociale: il latino era la lingua del potere, dell'amministrazione, dell'esercito, del commercio di lusso, della religione cristiana emergente. Chi voleva partecipare a questi sistemi di potere doveva imparare il latino; chi imparava il latino trasmetteva questa competenza ai propri figli come risorsa sociale preziosa. Il latino si diffuse per *elite dominance* — attraverso la dominanza di un gruppo piccolo ma potente — non per sostituzione demografica.

Questo caso è rilevante per la questione etrusca perché mostra che la diffusione di una lingua senza sostituzione demografica è un fenomeno storicamente documentato, che può avvenire su scala

regionale con tempistiche dell'ordine di secoli. Se il proto-etrusco si fosse diffuso nell'Italia centrale attraverso un meccanismo analogo a quello del latino in Britannia — cioè attraverso l'adozione da parte della popolazione locale di una lingua associata al prestigio commerciale o religioso di un gruppo piccolo ma influente — la firma genomica di questo evento potrebbe essere molto debole o addirittura assente, compatibilmente con ciò che la paleogenomica ha effettivamente osservato.

P.3 Il caso dei Proto-Indoeuropei in India: dibattito ancora aperto

Il caso della diffusione delle lingue indoeuropee nel subcontinente indiano — e in particolare la questione di quando e come il Sanscrito vedico sia arrivato in India — è forse il parallelo metodologico più complesso e più istruttivo per la questione etrusca, perché mostra come una controversia strutturalmente analoga (lingua non locale, continuità biologica apparente, ipotesi migratoria vs. autoctona) possa essere affrontata con gli stessi strumenti della paleogenomica, producendo risultati parzialmente analoghi ma in un contesto ancora più complicato.

Le lingue indoeuropee dell'India — il Sanscrito, i Prakrit, le lingue moderne derivate come il Bengali, l'Hindi, il Marathi, il Punjabi — sono chiaramente appartenenti alla famiglia indoeuropea (ramo indo-iranico), ma la loro origine geografica è stata a lungo controversa. L'ipotesi dominante nella linguistica comparativa — l'*Out of the Steppe* hypothesis — postulava che le lingue indoeuropee fossero arrivate in India attraverso la migrazione di popolazioni delle steppe pontico-caspie attraverso l'Afghanistan e il Pakistan, durante il II millennio avanti Cristo, in concomitanza con la crisi della Civiltà dell'Indo. Le tradizioni nazionaliste indiane — sia nella loro versione colonialista (la «invasione ariana») sia nella loro reazione nazionalista postcoloniale (la «Out of India» hypothesis) — avevano politicizzato fortemente questo dibattito, rendendo difficile una discussione scientifica serena.

Gli studi paleogenomici degli anni 2018–2021 — in particolare lo studio di Narasimhan e collaboratori su *Science* nel 2019 su duecentotrentasette individui dell'Asia centrale e meridionale dell'età del Bronzo e del Ferro — hanno mostrato risultati che sono, nel loro schema generale, paralleli a quelli etruschi ma in direzione opposta: hanno confermato la presenza di una componente significativa di ascendenza delle steppe (riconducibile alle culture Andronovo e BMAC, Bronze Age della Asia Centrale) nelle popolazioni del subcontinente indiano del II millennio avanti Cristo, compatibile con una migrazione demografica dall'Asia centrale verso il subcontinente. Questo risultato supporta l'ipotesi migratoria e contro l'ipotesi autoctona per la diffusione delle lingue indoeuropee in India — l'opposto di quanto accade per gli Etruschi.

Il confronto tra i due casi — India e Etruria — illustra la potenza discriminante della paleogenomica: in un caso (l'India) la firma di una migrazione demografica è presente e chiara; nell'altro (l'Etruria) è assente. Le stesse domande poste con gli stessi strumenti producono risposte diverse, perché la storia è diversa. Questo è il comportamento atteso di uno strumento scientificamente valido: produce risultati empiricamente determinati, non risultati predeterminati dalla teoria o dall'ideologia.

P.4 Il caso etnoarcheologico dei Baschi: la lingua sopravvissuta

Il Paese Basco — o più precisamente, le province basche di Spagna e Francia in cui il Basco è ancora parlato come lingua viva — offre il parallelo più diretto e più istruttivo per comprendere le implicazioni del risultato genomico etrusco sulla questione linguistica. Il Basco (*Euskara*) è una lingua pre-indoeuropea — l'unica sopravvissuta in Europa occidentale — completamente priva di parentele documentabili con qualsiasi altra lingua viva o morta del mondo (con la parziale eccezione delle lingue aquitane antiche, attestate in iscrizioni latine dei Pirenei e chiaramente imparentate con il Basco). I Baschi moderni hanno un profilo genomico che riflette la stessa struttura dell'ascendenza europea — neolitico anatolico + WHG + una proporzione di steppe inferiore alla media europea — e non mostrano nessuna componente di ascendenza «esotica» che potrebbe spiegare l'isolamento linguistico del Basco come effetto di una migrazione da un'area non indoeuropea.

Come è sopravvissuta la lingua basca attraverso la romanizzazione, le invasioni visigotica e moresche, la Reconquista, e secoli di pressione culturale e linguistica da parte del castigliano e del francese? La risposta è complessa ma identifica almeno tre fattori principali. Il **primo** è la geografia: i Pirenei formano una barriera naturale che ha rallentato la penetrazione delle lingue di prestigio circostanti e ha protetto il Basco nelle sue valli più isolate. Il **secondo** è la densità demografica: la popolazione basca nel suo nucleo originario aveva una densità sufficiente da resistere alla sostituzione linguistica per assimilazione, anche quando la lingua di prestigio circostante (prima il latino, poi il castigliano) si espandeva. Il **terzo** è l'identità culturale: il Basco è diventato nel corso dei secoli non soltanto una lingua ma un marcatore di identità etnica e culturale, la cui conservazione era percepita dalla comunità parlante come un valore in sé che giustificava lo sforzo di mantenerla viva contro la pressione delle lingue di maggioranza.

Tutti e tre questi fattori — barriere geografiche, densità demografica, identità culturale — sono in linea di principio applicabili alla situazione dell'Etruria nel I millennio avanti Cristo. L'Appennino formava una barriera relativa tra l'Etruria e le aree di penetrazione dei popoli italici; le città etrusche avevano le densità demografiche più alte dell'Italia pre-romana; l'identità culturale etrusca — con la

sua scrittura, la sua tradizione religiosa, il suo sistema istituzionale — era abbastanza coesa e abbastanza rispettata (anche a Roma, che adottò molte sue istituzioni) da poter svolgere il ruolo di fattore di coesione linguistica. Il parallelismo con il caso basco non dimostra che il proto-etrusco fosse un substrato locale sopravvissuto — è compatibile anche con gli Scenari B e C del Capitolo IV — ma dimostra che la sopravvivenza di una lingua pre-indoeuropea in una nicchia geografica specifica, con una popolazione biologicamente indistinguibile dai popoli indoeuropei circostanti, è un fenomeno che si è effettivamente verificato in Europa e non richiede spiegazioni straordinarie.

NOTA FINALE

Questa monografia è stata concepita come uno strumento di lavoro per lo studioso che voglia orientarsi nella letteratura sulla questione etrusca con gli strumenti dell'epistemologia contemporanea. Non è, né vuole essere, un'opera definitiva: la ricerca sulle origini etrusche continua, e gli anni a venire porteranno nuovi campioni paleogenomici, nuove analisi linguistiche computazionali, nuovi dati isotopici che affineranno ulteriormente il quadro. La solidità della conclusione principale — la continuità biologica tra Etruschi classici e Villanoviani dell'Italia del Bronzo Finale — non sarà probabilmente messa in discussione da questi nuovi dati; ma le questioni aperte che questa conclusione ha reso più nitide riceveranno risposte progressive che meritano di essere seguite con attenzione.

La lingua etrusca rimane il grande enigma irrisolto. Ogni nuovo testo etrusco di qualche lunghezza — ogni nuova lamina di bronzo o di piombo, ogni nuova parete con iscrizione dipinta — è un evento di straordinario valore per la comprensione della civiltà. Il corpus, per quanto già grande in termini numerici (oltre tredicimila testi), è troppo formulare per permettere la ricostruzione lessicale necessaria a valutare con strumenti formali moderni il grado di parentela con il lemnico. Una fortuna straordinaria — il ritrovamento di un altro testo bilingue lungo come le Laminette di Pyrgi ma in etrusco e in una lingua indoeuropea conosciuta — cambierebbe radicalmente questa situazione. È questo il sogno di ogni etruscologo: un testo lungo, bilingue, con un contenuto non formulare che permetta di ancorare il sistema linguistico etrusco al mondo delle lingue note. Finché questo testo non si trova — o non si riconosce tra i materiali già scavati e non ancora analizzati — l'etrusco rimarrà una lingua nota nella sua struttura grammaticale e opaca nel suo lessico. Ed è questa opacità, più di qualsiasi altra cosa, che alimenta il fascino perenne della civiltà che l'ha parlata.

La paleogenomica ha trasformato molte discipline delle scienze umane, ma il suo impatto sull'etruscologia è stato particolarmente profondo perché ha risposto in modo così diretto e così netto a una domanda che era rimasta aperta per venticinque secoli. Questo tipo di risposta — rapida, quantitativa, non negoziabile nel senso che i numeri non possono essere reinterpretati indefinitamente come le parole delle fonti letterarie — ha avuto un effetto psicologico e istituzionale sulla disciplina che è difficile sottovalutare. Gli etruscologi che avevano dedicato decenni alla questione delle origini si sono trovati di fronte a un risultato che rendeva le loro argomentazioni tradizionali — filologiche, archeologiche, craniometriche — secondarie rispetto a un dato di un tipo completamente nuovo. Alcuni hanno accolto il cambiamento con entusiasmo, riconoscendo che la genomica confermava in modo quantitativo conclusioni che la tradizione italiana dell'etruscologia — da Pallottino in poi — aveva già intuito in modo qualitativo. Altri hanno espresso riserve

metodologiche più o meno fondate. Ma la direzione del campo è ormai chiara: l'etruscologia del XXI secolo sarà una disciplina genuinamente interdisciplinare, che integra le competenze filologiche e archeologiche tradizionali con quelle della bioinformatica, della genetica delle popolazioni, dell'analisi isotopica e della modellistica demografica. Questa integrazione è la principale eredità metodologica della risoluzione genomica della questione etrusca, e il suo valore va ben oltre il caso specifico degli Etruschi: offre un modello per come le discipline umanistiche possano rinnovarsi attraverso il dialogo con le scienze naturali senza perdere la propria specificità e il proprio contributo insostituibile alla comprensione del passato umano.

Il lettore che si avvicina per la prima volta alla civiltà etrusca attraverso queste pagine troverà probabilmente che la domanda più interessante non è quella a cui la genomica ha risposto — da dove venivano biologicamente gli Etruschi — ma quella che la genomica ha reso più nitida senza rispondervi: come mai una popolazione biologicamente indistinguibile dai Latini, dagli Umbri e dagli altri popoli dell'Italia centrale abbia sviluppato e mantenuto per oltre settecento anni una lingua radicalmente diversa da tutte le lingue indoeuropee circostanti, un sistema di scrittura proprio, un patrimonio religioso e rituale originale, e una tradizione artistica di straordinaria potenza espressiva? La risposta a questa domanda è ancora da scrivere, e sarà la grande avventura intellettuale dell'etruscologia del XXI secolo.

Una parola finale sui limiti di questa monografia. Ho cercato di essere onesto nelle incertezze, esplicito nelle strutture logiche degli argomenti, e preciso nelle distinzioni tra ciò che è certo, ciò che è probabile e ciò che rimane aperto. Ho cercato di evitare la tentazione — sempre presente nella scrittura per un pubblico colto ma non specializzato — di presentare come risolto ciò che è soltanto avanzato, o come certo ciò che è soltanto altamente probabile. Se in qualche punto ho fallito in questo obiettivo, la colpa è dell'entusiasmo per una materia che trovo genuinamente affascinante nelle sue mille dimensioni — linguistica, archeologica, genetica, storica, filosofica. Gli Etruschi meritano tutto questo entusiasmo, e molto di più.

BIBLIOGRAFIA RAGIONATA

I. Fonti primarie

- Erodoto. Storie, I, 94. Edizione critica: C. Hude, *Herodoti Historiae*, 2 voll., Oxford: Clarendon Press, 1927.
Commento: W.W. How e J. Wells, *A Commentary on Herodotus*, 2 voll., Oxford 1912. Traduzione italiana: A. Fraschetti (a cura di), *Erodoto. Le Storie*, Milano: Mondadori, 1988. Analisi del passo specifico come aition eziologico: D.L. Page, *The Lydiaca of Herodotus*, Cambridge 1982.
- Dionisio di Alicarnasso. Antichità Romane, I, 26–30. Edizione Loeb: E. Cary, *Dionysius of Halicarnassus. Roman Antiquities*, 7 voll., Cambridge MA: Harvard University Press, 1937–1950. Studio del metodo storiografico: E. Gabba, *Dionysius and the History of Archaic Rome*, Berkeley: University of California Press, 1991.
- Tito Livio. Ab Urbe Condita, V, 33. Edizione critica: R.M. Ogilvie, *Titi Livi ab Urbe Condita libri I–V*, Oxford: Clarendon Press, 1974. Critica dell'interpretazione nordica: D. Briquel, *L'origine lydienne des Étrusques*, Roma 1991, pp. 34–41.
- Virgilio. Eneide, VIII, 479–495; X, 163–165. Edizione critica: R.A.B. Mynors, *P. Vergili Maronis Opera*, Oxford 1969. Commento: R.D. Williams, *The Aeneid of Virgil, Books 7–12*, London 1973.
- Plinio il Vecchio. *Naturalis Historia*, III, 50–51. Edizione Loeb: H. Rackham, *Pliny. Natural History*, 10 voll., Cambridge MA 1938–1963.
- Strabone. Geografia, V, 2, 1–9. Edizione Loeb: H.L. Jones, *The Geography of Strabo*, 8 voll., Cambridge MA 1917–1932.
- Varrone. *De Lingua Latina*, IV, 30–34. Edizione con commento: R.G. Kent, *Varro on the Latin Language*, 2 voll., Cambridge MA 1938.
- Cicerone. *De Divinatione*, I, 41–42; II, 51–52. Edizione con commento: W.A. Falconer, *Cicero. De Senectute, De Amicitia, De Divinatione*, Cambridge MA 1923.

II. Epigrafia e linguistica etrusca

- Agostiniani, L., e Nicosia, F. *Tabula Cortonensis*. Roma: L'Erma di Bretschneider, 2000. Edizione critica con trascrizione, traduzione e analisi linguistica.
- Agostiniani, L. «L'etrusco come lingua SOV: dati e problemi.» *Studi Etruschi* 54 (1986–87): 147–174. Studio fondamentale sull'ordine dei costituenti nella frase etrusca.
- Bonfante, G., e Bonfante, L. *The Etruscan Language: An Introduction*. Manchester: Manchester University Press, 1983; II ed. riveduta 2002. Il manuale più completo in lingua inglese su fonologia, morfologia e lessico.
- Bonfante, L. (a cura di). *Etruscan Life and Afterlife: A Handbook of Etruscan Studies*. Detroit: Wayne State University Press, 1986. Manuale di riferimento anglofono con contributi specialistici.

- Cristofani, M. Introduzione allo studio dell'etrusco. Firenze: Olschki, 1973; II ed. 1991. La migliore introduzione alla paleografia e all'epigrafia etrusca in italiano.
- De Simone, C. «Etrusco e Lemnico: il rapporto.» In *Studia Linguistica et Orientalia Memoriae Haim Rosén Dedicata*, Leuven 2001, pp. 139–165. Analisi più rigorosa del confronto lemnico-etrusco con documentazione sistematica delle isoglosse.
- De Simone, C. «La nuova iscrizione 'tirrenia' di Lemnos (Efestia, teatro): considerazioni generali.» *Rasenna: Journal of the Center for Etruscan Studies* 1/1 (2009). Discussione dell'iscrizione più recente scoperta a Lemno.
- Facchetti, G.M. L'enigma svelato della lingua etrusca. Firenze: Passigli, 2000. Sintesi accessibile con discussione critica dei metodi di decifrazione.
- Facchetti, G.M. Appunti di morfologia etrusca. Firenze: Olschki, 2002. Studio tecnico della morfologia verbale e nominale.
- Lejeune, M. «Les épicleses de l'inscription de Kaminia.» *Revue des Études Grecques* 79 (1966): 1–17. Analisi delle formule onomastiche nella Stele di Kaminia.
- Maggiani, A. «Le iscrizioni della Stele di Lemno.» *Studi Etruschi* 52 (1984): 115–143. Analisi paleografica comparata dell'alfabeto delle iscrizioni lemnice con quello etrusco arcaico.
- Marchesini, S. Le lingue frammentarie dell'Italia antica. Milano: Hoepli, 2009. Quadro comparativo sulla posizione dell'etrusco nel panorama delle lingue preromane.
- Pallottino, M. *Testimonia Linguae Etruscae*. Firenze: La Nuova Italia, 1954; II ed. 1968. Corpus epigrafico etrusco di riferimento per la generazione precedente il *Thesaurus*.
- Pallottino, M. *Saggi di antichità*. 3 voll. Roma: G. Bretschneider, 1979. Raccolta degli studi fondamentali di Pallottino su lingua, epigrafia e storia etrusca.
- Prosdocimi, A.L. «Osservazioni sull'etrusco.» *Studi Etruschi* 46 (1978): 199–222. Saggio metodologico sulla posizione dell'etrusco rispetto alle lingue indoeuropee dell'Italia antica.
- Rix, H. (a cura di). *Etruskische Texte: Editio minor*. Tübingen: Narr, 1991. 2 voll. L'edizione di riferimento del corpus con classificazione tipologica e geografica.
- Rix, H. «Tirsensisch, Etruskisch, Lemnisch.» In N. Oettinger e Th. Steer (a cura di), *Das Nomen im Indogermanischen*, Wiesbaden 2014, pp. 338–348.
- Steinbauer, D.H. *Neues Handbuch des Etruskischen*. St. Katharinen: Scripta Mercaturae, 1999. Il dizionario più aggiornato dell'etrusco con analisi morfologica sistematica.
- Thesaurus Linguae Etruscae*. Roma: CNR, 1978–. Opera in corso. Corpus più sistematico e aggiornato, consultabile online.

III. Archeologia ed etruscologia

- Barker, G., e Rasmussen, T. *The Etruscans*. Oxford: Blackwell, 1998; II ed. 2000. Il manuale anglofono più equilibrato sulla civiltà etrusca nella sua interezza.

- Bartoloni, G. *Le società dell'Italia primitiva: lo studio delle necropoli e la nascita delle aristocrazie*. Roma: Carocci, 2003. Studio fondamentale sulla struttura sociale delle comunità italiane del Bronzo Finale.
- Bellelli, V. (a cura di). *Le origini degli Etruschi: storia, archeologia, antropologia*. Roma: L'Erma di Bretschneider, 2012. Raccolta di contributi interdisciplinari sulla questione delle origini nello stato della ricerca pre-genomica.
- Briquel, D. *L'origine lydienne des Étrusques: histoire de la doctrine dans l'Antiquité*. Roma: École Française de Rome, 1991. Il contributo più completo alla storia dell'ipotesi anatolica dall'antichità all'età moderna.
- Briquel, D. *Les Pélasges en Italie: recherches sur l'histoire de la légende*. Roma: École Française de Rome, 1984.
- Camporeale, G. *Gli Etruschi: storia e civiltà*. Torino: UTET, 2000; III ed. 2011. La sintesi enciclopedia più completa sulla civiltà etrusca.
- Colonna, G. (a cura di). *Santuari d'Etruria*. Milano: Electa, 1985. Catalogo sistematico dei principali santuari etruschi.
- Cornell, T.J. *The Beginnings of Rome*. London: Routledge, 1995. Il contesto storico-politico più affidabile per i rapporti tra Etruria e Roma nel periodo arcaico.
- Cristofani, M. *L'arte degli Etruschi: produzione e consumo*. Torino: Einaudi, 1978. Studio fondamentale sulla produzione artistica etrusca come fenomeno economico e sociale.
- Cristofani, M. (a cura di). *Civiltà degli Etruschi*. Milano: Electa, 1985. Catalogo della grande mostra fiorentina.
- Haynes, S. *Etruscan Civilization: A Cultural History*. London: British Museum Press, 2000. Sintesi culturale equilibrata con ottima documentazione iconografica.
- Pallottino, M. *Etruscologia*. Milano: Hoepli, 1942; VII ed. 1984. Il manuale fondamentale della disciplina.
- Pallottino, M. «Le origini degli Etruschi: problemi e prospettive di ricerca.» *Dialoghi di Archeologia* 1 (1967): 5–22. La formulazione più chiara della posizione di Pallottino sulla distinzione tra domanda biologica e domanda culturale.
- Potter, T.W. *The Changing Landscape of South Etruria*. London: Elek, 1979.
- Rasmussen, T.B. *Bucchero Pottery from Southern Etruria*. Cambridge: Cambridge University Press, 1979.
- Sprenger, M., e Bartoloni, G. *Die Etrusker: Kunst und Geschichte*. München: Hirmer, 1977.
- Torelli, M. *Storia degli Etruschi*. Bari: Laterza, 1981; nuova ed. 1997. La sintesi storica più lucida dal punto di vista della storia politica e sociale.
- Turfa, J.M. (a cura di). *The Etruscan World*. London/New York: Routledge, 2013; aggiornamento 2022. La raccolta più aggiornata di contributi specialistici su tutti gli aspetti della civiltà etrusca.

IV. Paleogenomica: studi fondamentali e comparativi

- Achilli, A., et al. «Mitochondrial DNA variation of modern Tuscans supports the Near Eastern origin of Etruscans.» *American Journal of Human Genetics* 80 (2007): 759–768. Studio mtDNA precursore; conclusioni oggi superate ma metodologicamente istruttivo.
- Alexander, D.H., Novembre, J., e Lange, K. «Fast model-based estimation of ancestry in unrelated individuals.» *Genome Research* 19 (2009): 1655–1664. L'articolo che introduce ADMIXTURE.
- Allentoft, M.E., et al. «Population genomics of Bronze Age Eurasia.» *Nature* 522 (2015): 167–172. Studio parallelo a Haak et al. 2015 sulle migrazioni delle steppe.
- Brandt, G., et al. «Ancient DNA Reveals Key Stages in the Formation of Central European Population Structure.» *Science* 342 (2013): 257–261.
- Fu, Q., et al. «An early modern human from Romania with a recent Neanderthal ancestor.» *Nature* 524 (2015): 216–219. Rilevante per i metodi di autenticazione del DNA antico.
- Haak, W., et al. «Massive migration from the steppe was a source for Indo-European languages in Europe.» *Nature* 522 (2015): 207–211. Lo studio fondativo della paleogenomica europea su larga scala; benchmark per la scala delle migrazioni delle steppe.
- Ioannidis, A.G., et al. «Native American gene flow into Polynesia predating Easter Island settlement.» *Nature* 583 (2020): 572–577. Il caso polinesiano-sudamericano: contatto genomico pre-colombiano rilevato a piccola scala.
- Lazaridis, I., et al. «Ancient human genomes suggest three ancestral populations for present-day Europeans.» *Nature* 513 (2014): 409–413. Introduce il modello delle tre componenti ancestrali EEF/WHG/Steppe.
- Lazaridis, I., et al. «Genomic insights into the origin of farming in the ancient Near East.» *Nature* 536 (2016): 419–424. Caratterizza le popolazioni sorgente della componente neolitico-anatolica degli Europei.
- Mathieson, I., et al. «Genome-wide patterns of selection in 230 ancient Eurasians.» *Nature* 528 (2015): 499–503.
- Narasimhan, V.M., et al. «The Formation of Human Populations in South and Central Asia.» *Science* 365 (2019): eaat7487. Il caso dei Proto-Indoeuropei in India: migrazione demografica delle steppe verso il subcontinente confermata dalla paleogenomica.
- Olalde, I., et al. «The Beaker phenomenon and the genomic transformation of northwest Europe.» *Nature* 555 (2018): 190–196. Mostra la quasi completa sostituzione demografica dell'Inghilterra neolitica; benchmark fondamentale per la scala di una migrazione massiccia.
- Patterson, N., et al. «Ancient admixture in human history.» *Genetics* 192 (2012): 1065–1093. Introduce le statistiche-f e i metodi qpAdm/qpGraph; fondamentale per tutta l'analisi del Cap. II.
- Posth, C., et al. «The origin and legacy of the Etruscans through a 2000-year archeogenomic time transect.» *Science Advances* 7 (2021): eabi7673. Lo studio paleogenomico principale sugli Etruschi. Open access.

- Reich, D. *Who We Are and How We Got Here*. New York: Pantheon, 2018; trad. it. *Come siamo arrivati fin qui*, Milano: Raffaello Cortina, 2019. La presentazione divulgativa più autorevole della paleogenomica.
- Rohland, N., et al. «Partial uracil-DNA-glycosylase treatment for screening of ancient DNA.» *Philosophical Transactions of the Royal Society B* 370 (2015): 20130624. Protocollo di autenticazione del DNA antico.
- Schiffels, S., et al. «Iron Age and Anglo-Saxon genomes from East England reveal British migration history.» *Nature Communications* 7 (2016): 10408. Il caso britannico post-romano: continuità biologica celtica nonostante la diffusione del latino e dell'anglosassone.
- Skoglund, P., e Mathieson, I. «Ancient Genomics of Modern Humans: The First Decade.» *Annual Review of Genomics and Human Genetics* 19 (2018): 381–404. Rassegna metodologica aggiornata.
- Villalba-Mouco, V., et al. «Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula.» *Current Biology* 29 (2019): 1169–1177.

V. Linguistica computazionale, comparata e glottocronologia

- Atkinson, Q.D., e Gray, R.D. «How old is the Indo-European language family?» In P. Forster e C. Renfrew (a cura di), *Phylogenetic Methods and the Prehistory of Languages*, Cambridge 2006, pp. 91–109.
- Baxter, W.H., e Manaster Ramer, A. «Beyond lumping and splitting: probabilistic issues in historical linguistics.» In C. Renfrew et al. (a cura di), *Time Depth in Historical Linguistics*, Cambridge 2000, vol. I, pp. 167–188. Il test statistico per la probabilità di coincidenza casuale tra isoglosse.
- Chang, W., et al. «Bayesian phylogenetics using an augmented alignment model.» *Language* 91 (2015): 194–209. Il modello bayesiano per la filogenetica linguistica discusso nella sezione IV.2.4.
- Fortson, B.W. *Indo-European Language and Culture: An Introduction*. Oxford: Blackwell, 2004; II ed. 2010.
- Gray, R.D., e Atkinson, Q.D. «Language-tree divergence times support the Anatolian theory of Indo-European origin.» *Nature* 426 (2003): 435–439. Il metodo della glottocronologia bayesiana applicato alle lingue indoeuropee.
- Ringe, D., e Warnow, T. «Computational cladistics and the position of Tocharian.» In W. Meid (a cura di), *Sprache und Kultur der Indogermanen*, Innsbruck 1998, pp. 391–414. Il test formale di parentela linguistica citato nella sezione IV.5.2.
- Swadesh, M. «Lexicostatistic dating of prehistoric ethnic contacts.» *Proceedings of the American Philosophical Society* 96 (1952): 452–463. Il metodo originale della glottocronologia, discusso criticamente nella sezione IV.2.4.
- Trask, R.L. *The History of Basque*. London: Routledge, 1997. Lo studio linguistico più completo sulla storia del basco e sul suo isolamento genetico: parallelo principale per lo Scenario A.
- Trask, R.L. «Where do the Basques come from?» *Fontes Linguae Vasconum* 31 (1999): 5–45. Sintesi sulle teorie dell'origine del basco.

VI. Epistemologia, storia della storiografia e sociologia della scienza

- Bayes, T. «An Essay towards Solving a Problem in the Doctrine of Chances.» *Philosophical Transactions of the Royal Society* 53 (1763): 370–418. La formulazione originale del teorema di Bayes.
- Briquel, D. «Origine lydienne, origine alpine: deux visions des Étrusques.» *Revue études latines* 74 (1996): 21–38. Analisi storico-storiografica delle due grandi tradizioni.
- Ginzburg, C. «Spie: radici di un paradigma indiziario.» In A. Gargani (a cura di), *Crisi della ragione*, Torino: Einaudi, 1979, pp. 57–106. Il paradigma indiziario come metodo di inferenza storica.
- Hacking, I. *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press, 2001.
- Hacking, I. *The Emergence of Probability*. Cambridge: Cambridge University Press, 1975; II ed. 2006. Storia concettuale del pensiero probabilistico.
- Hessé, M.B. *Models and Analogies in Science*. Notre Dame: University of Notre Dame Press, 1966. Il quadro teorico per la struttura delle analogie scientifiche: citato nella sezione IV.3.5.
- Kahneman, D. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux, 2011; trad. it. *Pensieri lenti e veloci*, Milano: Mondadori, 2012. La sintesi più accessibile degli effetti di ancoraggio cognitivo.
- Kahneman, D., e Tversky, A. «Judgment under uncertainty: heuristics and biases.» *Science* 185 (1974): 1124–1131. L'articolo originale sull'effetto di ancoraggio cognitivo: citato nella sezione IV.3.2.
- Kahan, D.M., et al. «The polarizing impact of science literacy and numeracy on perceived climate change risks.» *Nature Climate Change* 2 (2012): 732–735. Il motivated reasoning e il cultural cognition: rilevante per la sezione III.
- Kahan, D.M. «Misconceptions, misinformation, and the logic of identity-protective cognition.» *Cultural Cognition Project Working Paper Series* 164 (2017).
- Kuhn, T.S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962; IV ed. 2012; trad. it. *La struttura delle rivoluzioni scientifiche*, Torino: Einaudi, 1999.
- Lakatos, I. *The Methodology of Scientific Research Programmes*. Cambridge: Cambridge University Press, 1978. Il concetto di programma di ricerca scientifico come alternativa critica a Kuhn.
- Mastrocinque, A. «Ricerche sulle origini degli Etruschi: storia delle dottrine.» *Memorie dell'Accademia delle Scienze di Torino, Classe di Scienze Morali, Storiche e Filologiche*, serie V, 5 (1981): 1–133. La storia storiografica più completa della questione delle origini.
- Poliakov, L. *Le mythe aryen*. Paris: Calmann-Lévy, 1971; trad. it. *Il mito ariano*, Milano: Einaudi, 1976. Per il contesto ideologico in cui l'ipotesi alpina ha trovato trazione nel XIX–XX secolo.
- Popper, K.R. *The Logic of Scientific Discovery*. London: Hutchinson, 1959; trad. it. *Logica della scoperta scientifica*, Torino: Einaudi, 1970.
- Shapin, S. *A Social History of Truth*. Chicago: University of Chicago Press, 1994. La sociologia della conoscenza scientifica e la costruzione dell'autorità epistemica.
- Thagard, P. «The best explanation: Criteria for theory choice.» *Journal of Philosophy* 75 (1978): 76–92.